

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

FONDEMENTS ET DIVERSITÉ DE LA PHILOSOPHIE NÉO-MÉCANISTE DES
SCIENCES

MÉMOIRE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN PHILOSOPHIE

PAR
DAVY MOUGENOT

MARS 2006

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 -Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

Remerciements

Ce mémoire doit son existence, en premier lieu, à Luc Faucher. Il a eu la brillante idée de réorienter mes intérêts réductionnistes en me pointant l'existence du courant néo-mécaniste en philosophie des sciences. Je m'en rends compte à présent, l'organe actif du réductionnisme est l'ensemble des méthodes expérimentales, pas la philosophie, ou alors, peut-être, la philosophie des sciences.

Le mécanisme est devenu plus qu'un système philosophique, au sens limité d'un ensemble organisé de propositions, mais une perspective philosophique. Avoir travaillé sur ce sujet m'a permis d'imaginer le monde autrement, à la fois plus simplement et de façon plus riche. Plus simplement parce que l'intuition fondamentale du mécanisme est que tout est la conséquence dynamique de deux entités : les choses et les forces. De façon plus riche car j'ai aperçu la complexité qui se cache sous l'apparente simplicité de ces deux mots.

La diversité matérielle ainsi que son organisation, pour l'étudiant en philosophie que j'étais, ne m'est jamais apparue comme un fait digne d'intérêt : trop de philosophie classique, probablement. La fréquentation des neurosciences m'a amené à reconsidérer radicalement ce point de vue : le virus réductionniste m'avait contaminé. Ce pourrait-il que la matière, sous ce terme anodin, soit responsable de tout ?

Au grand dam de mes parents et de nombreux interlocuteurs, je me suis assez complètement dévoué à la connaissance, et la défense, de ce point de vue. Je me plaisais à imaginer les mécanismes sous-jacents des phénomènes, surtout ceux permettant des capacités aussi étonnantes que l'esprit.

Les résistances culturelles liées à la mécanisation du monde sont assez surprenantes et, au gré des conversations, à ma grande surprise, une forme de cartésianisme n'est pas en voie de disparition. Des exemples contemporains, surtout en sociologie de la communication (Lafontaine 2003), associent la pensée mécaniste à un anti-humanisme. Selon eux, la dignité humaine repose sur une certaine impénétrabilité de la conscience : toute tentative scientifique de compréhension est vue comme un viol de la personnalité. L'avenir, il me semble, est sombre pour ce genre de sympathies.

Sans vouloir verser dans la phénoménologie du cyborg et dans une vision haeckelienne de la politique, il m'apparaît indiscutable qu'une modification des conditions

matérielles du corps humain entraînera des conséquences substantielles sur la façon dont nous nous percevons et la manière dont nous nous gouvernerons (Fukuyama 2004).

Mes pensées vont, inévitablement, à Annie : je me demande parfois si l'ensemble de ce qui compose ma personne n'est pas réparti sur deux corps physiques (et quelque peu sur le disque dur de mon ordinateur). À un point tel que je devrais, peut-être, reconnaître une certaine forme d'externalisme : cela fera plaisir à un ami éloigné, mais inestimable.

Je me dois également de remercier aussi les professeurs Pierre Poirier et Alain Voizard qui m'ont soutenu, autant intellectuellement que pour les fastidieuses démarches administratives des demandes de bourses.

Table des matières

REMERCIEMENTS	i
TABLE DES MATIÈRES	iii
FIGURES	v
RÉSUMÉ	vi
INTRODUCTION	1
CHAPITRE 1	9
<i>La mise en place du néo-mécanisme</i>	
I. Théorie hempelienne de l'explication	13
II. Le tournant réaliste (Salmon)	18
1) Le durcissement des positions	19
2) La philosophie causale-mécanique	23
Le fantôme écossais	24
Ghostbuster	26
3) L'explication causale-mécaniste	32
III. Le tournant naturaliste et pragmatiste (Simon et Wimsatt)	35
1) La nouvelle ontologie	38
2) Les arguments en faveur de la nouvelle ontologie	45
CHAPITRE 2	
<i>Première théorie exhaustive du néo-mécanisme (Glennan/Bechtel/Richardson)</i>	52
I. Analyse du concept de mécanisme (Glennan)	55
1) L'ontologie mécaniste	57
Partie	58
Interaction directe	60
2) L'épistémologie mécaniste	62
Les modèles mécanistes	63
Théorie de la similarité	65

II. L'explication mécaniste comme stratégie de recherche scientifique (Bechtel et Richardson)	68
1) L'explication mécaniste comme heuristique de décomposition et localisation	70
2) Heuristique, décomposabilité et système complexe	77
CHAPITRE 3	
<i>Seconde théorie exhaustive du néo-mécanisme</i> (Machamer/Darden/Craver)	89
I. Analyse ontologique du mécanisme	91
1) Continuité productive	93
2) Entité et activité	94
3) Organisation spatiotemporelle	99
II. Les caractéristiques épistémiques du mécanisme	104
1) Les schémas mécanistes	105
2) Les sketches mécanistes	108
3) Niveau et hiérarchie mécaniste	108
4) Les explications causales-mécanistes	112
5) Les stratégies et méthodes expérimentales	116
CONCLUSION	121
BIBLIOGRAPHIE	128

Figures

1) Le modèle D-N	15
2) Le principe de la transmission de marque	30
3) Théorie mécaniste de l'explication	32
4) Typologie des interactions fondamentales	33
5) L'immédiacité mécanistique	61
6) Taxonomie des systèmes et décomposabilité	80
7) Heuristique de la décomposition et de la localisation	82
8) Mécanisme de la transmission chimique interneuronale	92
9) Typologie des activités	98
10) Micelles et bicouches	103
11) Représentation multi-niveaux hiérarchisée de la mémoire spatiale	109
12) L'explication mécaniste	114
13) Expérimentations ascendantes (bottom up) et descendantes (top down)	117

Résumé

Aborder le mécanisme sans référer préalablement à Descartes est encore difficile, même de nos jours. Pourtant, le mécanisme a connu une constante évolution depuis le XIX^{ème} siècle et, surtout, il a ressurgi depuis plusieurs décennies dans nos sciences et notre philosophie des sciences. C'est pour essayer de palier, bien modestement, ce retard que ce mémoire est consacré au néo-mécanisme, la version contemporaine du mécanisme.

Il ne faut pas croire que cette résurgence est un simple fait historique : nous sommes d'avis que le néo-mécanisme apporte des positions articulées, cohérentes et originales concernant l'ontologie (réductionnisme non-éliminativiste) et l'épistémologie (théorie de l'explication scientifique).

Ce mémoire expose, dans le premier chapitre, comment les principes de base du néo-mécanisme, soit le réalisme et le naturalisme pragmatiste, se sont forgés en réaction aux thèses philosophiques des empiristes logiques concernant la théorie de l'explication scientifique et la nature de la causalité. Les chapitres deux et trois présentent, respectivement, deux théories néo-mécanistes exhaustives, c'est-à-dire une théorie qui couvrent les aspects ontologique, épistémologique et méthodologique.

Nous verrons, au chapitre deux, comment la première théorie exhaustive du néo-mécanisme, développée par Glennan, Bechtel et Richardson, s'est construite en continuité avec les projets ontologique et épistémologique de Simon et Wimsatt en traitant la notion de mécanisme sous une perspective systémique et les explications mécanistes sous une perspective méthodologique et heuristique (une stratégie de recherche scientifique). Cette première « vague » de néo-mécanistes affirme que le néo-mécanisme est la forme la plus répandue des explications scientifiques tant son application est diffusée.

Le troisième chapitre présente la seconde théorie exhaustive du néo-mécanisme (Machamer/Darden/Craver). Leur théorisation du concept de mécanisme et de l'explication mécaniste s'est faite en référence à la neurobiologie et la biologique moléculaire – la seconde « vague » de néo-mécanistes ne prétend donc pas que le néo-mécanisme soit une théorie générale de l'explication scientifique. Ces néo-mécanistes ouvriront les premiers débats sur des enjeux internes à la doctrine néo-mécaniste.

Introduction

La coquille d'un œuf de poule est faite d'une matière blanche et délicate, et c'est une machine au même titre que le coquetier qui est fait pour la recevoir : l'une et l'autre sont deux modes d'une même fonction. Sans doute la poule fait la coquille à l'intérieur de son corps, mais cela n'empêche pas que ce soit tout simplement de la poterie. Elle fait son nid en dehors de son corps parce que cela lui est plus commode, mais le nid est une machine ni plus ni moins que la coquille d'œuf. Ce que l'on nomme « machine » n'est qu'un expédient.

Samuel Butler, *Erewhon*, Le livre des machines, p. 237.

Le mécanisme, comme l'empirisme, est un phoenix philosophique : il réapparaît à plusieurs reprises sous une forme nouvelle au cours de l'histoire occidentale. Même si l'on peut qualifier certains philosophes de l'Antiquité de « mécanistes », les débuts officiels de la doctrine mécaniste reste le mécanisme cartésien. Descartes impose un mandat immense à cette doctrine puisqu'elle était supposée offrir une alternative métaphysique et épistémologique systématique à la doctrine scholastique d'obédience aristotélicienne. La « mécanisation du monde » (Dijksterhuis, 1961 et Boas, 1956) devait nous débarrasser de toutes les entités considérées avec suspicion, comme les qualités et la notion de cause finale (et formelle), pour faire place à une vision métaphorique où le fonctionnement de la nature ne se distingue pas de celui d'une machine.

Hormis l'esprit humain, toute entité ou phénomène *devait* être composé ultimement de particules dont la nature et l'agencement *devaient* permettre d'expliquer les propriétés des entités ou phénomènes en question. Le mécanisme cartésien était donc une doctrine métaphysique (atomisme) et une doctrine épistémologique (théorie de l'explication) d'une portée générale, puisqu'elle s'étendait de la microphysique à la biologie¹.

¹ Il est aussi possible de faire valoir un aspect idéologique du mécanisme en ce qu'il s'oppose à la tradition scholastique et propose une vision du monde systématique et cohérente.

Le mécanisme en physique entre le XVII^{ème} et le XVIII^{ème}-XIX^{ème} a été rapidement accepté – peu de scientifiques ou de philosophes s’y sont opposés. La mort du mécanisme cartésien, au XVIII^{ème}-XIX^{ème} est essentiellement due à la double introduction incompatible de l’action à distance et de l’électromagnétisme comme force scientifiquement légitime. En biologie, toutefois, le parcours est davantage sinueux étant donné les très fortes oppositions qui s’élèvent contre la vision mécaniste du vivant (essentiellement le vitalisme). C’est dans ce climat de tension théorique et idéologique qu’est née la version que nous nommons « classique » du mécanisme – représentée typiquement par Claude Bernard. Le principal problème du mécanisme appliqué au vivant réside dans son apparente incapacité à expliquer les propriétés spécifiques du vivant – surtout l’aspect manifestement téléologique (qui était souvent associé aux causes finales) de l’organisation corporelle. Au XIX^{ème}, le mécanisme appliqué à la biologie est si peu accepté que l’*Introduction à l’étude de la médecine expérimentale* est un vaste plaidoyer consacré à montrer qu’une étude purement matérialiste du vivant est possible – néanmoins une sensibilité à la complexité de l’organisation des mécanismes.

Il est important de saisir les thèses constitutives qui font du mécanisme, indépendamment de ses occurrences historiques particulières, une perspective philosophique. Il repose sur une thèse ontologique, le réalisme physicaliste, et une théorie de l’explication, le réductionnisme : le monde est composé d’un certain nombre d’entité et de certains types d’interaction, toute explication de n’importe quel phénomène doit pouvoir être construite à l’aide de ces deux catégories. Le mécanisme comporte trois avantages inestimables. D’abord, il fournit une norme d’intelligibilité de la nature, c’est-à-dire qu’il fonde ontologiquement la possibilité de connaître le monde étant donnée la connaissance des entités et des interactions. Ensuite, il est heuristique, il indique comment expliquer un phénomène : il faut identifier les entités constituantes et les interactions présentes afin de montrer comment la production du phénomène est effective. Finalement, le mécanisme est une entreprise unificatrice : tout objet doit être constitué des entités reconnues par le réalisme physicaliste et toute

explication doit être élaborée de façon à élucider le fonctionnement interne de l'explanandum. La question de la nature et de la diversité des mécanismes est une question empirique, sujette aux variations courantes du travail de recherche scientifique.

Un tel projet peut susciter du scepticisme. Historiquement, nous l'avons brièvement mentionné, l'approche mécaniste sous sa forme cartésienne est unanimement déclarée obsolète, sa théorie ontologique ne comprenant qu'un seul type d'interaction ou d'activité², la physique en reconnaissant actuellement quatre³. Ceci suggère une limite conceptuelle : le mécanisme est dépendant des sciences. Toute modification des croyances scientifiques au sujet de l'ontologie affecte le programme mécaniste. À ce propos, nous verrons comment la résurgence du néo-mécanisme est principalement due aux sciences biologiques des dernières décennies, particulièrement la biologie moléculaire et les neurosciences. D'ailleurs, l'influence de la biologie sur les sciences, et la philosophie des sciences, se fera particulièrement via la notion de complexité (et son traitement par la théorie des systèmes complexes). La légitimité accordée à ces « nouvelles » sciences aura donc profondément modifié les rapports entre les sciences et la philosophie : pour la première fois depuis la Renaissance, la physique n'est plus *le* modèle de légitimité scientifique et *le* modèle d'une théorie ontologique générale.

Une autre limite est, pour ainsi dire, géographique. De nombreux auteurs affichent un scepticisme concernant la portée du mécanisme : en microphysique, le flou entourant les notions d'entité et les difficultés conceptuelles liées au principe d'incertitude rendent l'identification et la localisation des entités difficile (Glennan, 1992 ; Woodward, 2002) ; dans les sciences de l'esprit, certains objectent que la

² La force mécanique.

³ Les forces mécanique, gravitationnelle, électromagnétique et nucléaire (faible et forte) (Davies, 1996). Notons toutefois qu'il existe des tentatives d'unification de ces diverses forces en des forces plus fondamentales. Des ponts entre les forces nucléaires faible, électromagnétique et mécanique sont déjà établis, alors que la force gravitationnelle n'est pas encore intégrée.

représentation n'est pas analysable mécanistiquement⁴ (Von Eckardt et Poland, 2005) et, en sciences sociales, de nombreux auteurs estiment que le mécanisme, sous le couvert de l'individualisme méthodologique, est fondamentalement déficient à expliquer les actions collectives d'agents intentionnels (Searle 2001 ; Mayntz, 2003) et d'autres dénoncent le mécanisme parce qu'il est « déshumanisant » (Lafontaine, 2003).

Ce mémoire est consacré au néo-mécanisme, la version contemporaine du mécanisme. Il expose, dans le premier chapitre, comment les principes de base du néo-mécanisme, soit le réalisme (W. Salmon) et le naturalisme pragmatiste (H. Simon et W. Wimsatt), se sont forgés en réaction aux thèses philosophiques des empiristes logiques (Hempel, Oppenheim et Nagel) concernant la théorie de l'explication scientifique et la nature de la causalité. Le premier chapitre est donc, premièrement, une mise en perspective de la résurgence du mécanisme en montrant comment ses « fondateurs » ont construit les bases théoriques du renouveau mécaniste en réaction au modèle déductif-nomologique de l'empirisme logique. D'une approche normative, formaliste et linguistique du contexte de justification des théories, les néo-mécanistes vont passer à une approche descriptive, informelle et cognitive du contexte de découverte et du contexte de justification. Ce chapitre montrera également comment les néo-mécanistes (Simon et Wimsatt, en premier lieu) se sont distancés d'une ontologie physicaliste réductionniste pour défendre plutôt un physicalisme non-élimativiste plus en accord avec l'état des sciences contemporaines – essentiellement, les approches systémiques modernes des sciences biologiques.

Les chapitres deux et trois présentent, respectivement, deux théories néo-mécanistes exhaustives, c'est-à-dire une théorie ontologique, épistémologique et méthodologique. Nous verrons, au chapitre deux, comment la première théorie exhaustive du néo-mécanisme, développée par Glennan, Bechtel et Richardson, s'est

⁴ Nous utilisons l'anglicisme « mécanistique » pour éviter « mécaniquement » qui jette une certaine ambiguïté. « Mécaniquement » peut référer au type de force de la nature, les forces mécaniques au sens d'action par contact (pousser, tirer, ...) alors que mécanistique réfère à la thèse mécaniste qui englobe tous les types de forces reconnus par les sciences actuelles.

construite en continuité avec les projets ontologiques et épistémologiques de Simon et Wimsatt en traitant la notion de mécanisme sous une perspective systémique et les explications mécanistes sous une perspective méthodologique et heuristique (une stratégie de recherche scientifique). La position de Glennan s'inscrit dans une perspective de clarification conceptuelle des implications ontologiques et épistémologiques de la notion même de mécanisme. Il offre une théorie néo-mécaniste basée sur une approche systémique mais se concentre sur une analyse explicite du concept de mécanisme⁵. Sa théorie est supposée s'appliquer à tous les domaines scientifiques – exception faite de la microphysique. Cette thèse de Glennan n'est pas seulement symptomatique de l'importance des sciences biologiques dans la détermination de ce qu'est une explication scientifique : Glennan marginalise une partie importante de la physique en défendant que la microphysique n'est pas mécanistiquement analysable. Ainsi, l'ensemble des sciences utilise le mode explicatif mécaniste, sauf la microphysique. La physique ne représente donc plus le paradigme de l'explication scientifique mais plutôt une exception « régionale » dans la géographie du domaine scientifique.

Bechtel et Richardson, quant à eux, s'intéressent plus particulièrement aux explications mécanistes et, surtout, à la façon dont elles sont découvertes. Ils soutiennent que les explications mécanistes sont produites suite à un processus heuristique de localisation et de décomposition des systèmes complexes et pensent que leur approche est valide pour l'ensemble des sciences du vivant⁶. Nous constaterons que les deux projets développés parallèlement sont complémentaires et que, réunis, ils forment la première théorie exhaustive et générale du néo-mécanisme.

Le troisième chapitre présente la seconde théorie exhaustive du néo-mécanisme (Machamer/Darden/Craver : MDC). Leur théorisation du concept de

⁵ « A mechanism for a behavior is a complex system that produces that behavior by the interaction of a number of parts, where the interactions between parts can be characterized by direct, invariant, change-relating generalizations », in Glennan (2002) *Rethinking Mechanistic Explanation*, p. S344

⁶ De nombreux domaines sont couverts : génétique, biochimie, neurobiologie et les sciences cognitives.

mécanisme et de l'explication mécaniste s'est faite en référence à la neurobiologie et la biologie moléculaire – la seconde « vague » de néo-mécanistes ne prétend donc pas que le néo-mécanisme est une théorie générale de l'explication scientifique. En plus d'abandonner la conceptualisation systémique de la notion de mécanisme⁷, MDC reprocheront à la notion de mécanisme défendue par Glennan son inapplicabilité à leurs domaines de référence (biologie moléculaire et neurosciences). Ainsi, MDC ouvriront les premiers débats sur des enjeux internes à la doctrine néo-mécaniste. Sur le plan épistémologique, MDC soutiennent que les scientifiques tentent de rendre compte de la productivité constante des mécanismes par des schémas ou des sketchs mécanistes qui se présentent de façon privilégiée par des artefacts visuels mais aussi sous forme simplement linguistique. Méthodologiquement, nous verrons comment le néo-mécanisme s'appuie sur les stratégies et méthodes expérimentales des neurosciences (le mécanisme est toujours dépendant des méthodes scientifiques).

Notre catégorisation des positions néo-mécanistes selon leur conceptualisation de la notion de mécanisme et sur leur portée explicative est une façon commode de présenter le sujet. D'une part, la question générale de l'importance du mécanisme pour une théorie de l'explication scientifique dépend de l'extension fixée par la définition du mécanisme. Pour Glennan (1992), par exemple, la notion de mécanisme est tout à fait centrale pour une théorie de l'explication scientifique, de la physique à l'économie, car sa définition se veut trans-disciplinaire. À l'inverse, la définition du mécanisme de MDC n'a qu'une portée limitée puisqu'elle ne doit être évaluée que de façon interne aux neurosciences et à la biologie moléculaire⁸. Donc, la motivation d'aborder les multiples perspectives néo-mécanistes selon le point de vue de leur portée est due au rôle déterminant que joue celle-ci pour la question plus générale de l'importance du néo-mécanisme pour la philosophie des sciences (surtout sur

⁷ En le définissant comme « des entités et des activités organisées spatiotemporellement de conditions initiales à des conditions finales et sont producteurs de changements réguliers », in Machamer, Peter, Lindley Darden, and Carl Carver (2000), *Thinking About Mechanisms*, p. 3.

⁸ La question générale concernant l'applicabilité du néo-mécanisme à d'autres domaines, surtout la psychologie et la sociologie, ne sera pas abordée dans ce mémoire mais fera l'objet ma thèse de doctorat.

l'importance de la notion de mécanisme pour une théorie de l'explication scientifique). D'autre part, comprendre les enjeux et les problématiques inhérents au néo-mécanisme exige de connaître la première génération de néo-mécanistes avant la seconde puisque MDC critique plusieurs aspects de la première théorie du néo-mécanisme. Enfin, la portée octroyée à la doctrine mécaniste a toujours été, historiquement, un sujet de fiévreux débats : notamment, comme nous l'avons déjà mentionné, concernant la mécanisation du vivant au XIX^{ème} siècle et, de nos jours, la mécanisation de l'esprit et des sciences sociales.

Les problèmes ouverts par une approche mécaniste contemporaine en philosophie des sciences sont multiples. La première problématique sera celle de la réapparition même du mécanisme dans le contexte de la philosophie des sciences de fin de XX^{ème} siècle. En effet, nous verrons que la doctrine dominante en philosophie des sciences était l'empirisme logique qui est globalement aux antipodes de la philosophie mécaniste. Nous verrons donc, lors du premier chapitre, comment les fondements philosophiques du néo-mécanisme se sont opposés à ceux de l'empirisme logique.

La deuxième problématique, qui est centrale à ce mémoire, est la nécessité de caractériser de façon claire et explicite la notion même de mécanisme – tâche qui a été principalement menée par Glennan pour la première génération de néo-mécanistes (c'est pourquoi les débats s'effectuent surtout entre celui-ci et MDC). En effet, Glennan propose, pour la première fois dans l'histoire de la philosophie des sciences, une définition explicite et élaborée de ce concept (chapitre deux). Jusqu'à lors, la notion de mécanisme avait toujours été une notion intuitive, un « proto-concept ». Nous verrons que plusieurs notions incluses dans la notion de mécanisme poseront problème pour la seconde génération de néo-mécanistes, notamment le recours aux lois pour caractériser les relations entre les parties de mécanismes, mais, plus fondamentalement, nous aborderons les différends quant au statut ontologique des interrelations (sont-elles réductibles à des dispositions d'objets ou sont-elles des entités ontologiques distinctes des objets ?). Aussi, un autre problème ontologique se

posera quant à la question de savoir quelles entités sont (et doivent être) acceptables comme « composants » d'un mécanisme (chapitre trois).

La troisième problématique que nous aborderons sera donc celle de la découverte effective des mécanismes : si les phénomènes biologiques sont si complexes, comment parvenir à la découverte de ces mécanismes ? Nous avons vu que le mécanisme classique de Bernard a fait le pari que les organismes sont des ensembles hiérarchiquement organisés de déterminismes complexes (ou mécanismes) qui *devraient* être réductibles à une organisation complexe et organisée de mécanismes simples. La réponse néo-mécaniste, présentée par Bechtel et Richardson (à la suite du travail de Wimsatt), sera présentée dans la deuxième partie du chapitre deux : les mécanismes sont découverts en procédant, par une série d'étape successive, à la décomposition du système étudié et à la localisation des ses composants.

Le néo-mécanisme jette un regard nouveau sur les approches contemporaines en philosophie des sciences qui, après Kuhn, sont réticentes à suggérer des critères normatifs aux sciences. En effet, en statuant sur les entités acceptables comme constituant (ou partie) d'un mécanisme, les néo-mécanistes excluent nécessairement certaines entités de la théorie néo-mécaniste de l'explication scientifique. C'est donc dire que les néo-mécanistes ne se contentent pas de décrire les pratiques scientifiques mais énoncent des critères contraignants pour les sciences. Par exemple, Glennan (1992) exclut les entités de la microphysique du domaine des phénomènes mécanistiquement explicable – marquant clairement l'importance récente et majeure des sciences biologiques dans la géographie du monde scientifique.

Finalement, ce mémoire se veut une présentation problématisée de la récente résurgence du mécanisme en philosophie des sciences. Nous comptons montrer que le néo-mécanisme propose au moins deux théories exhaustives de l'explication scientifique. À ce titre, ce mémoire fera valoir que le néo-mécanisme est certainement une des théories de l'explication scientifique la plus complète actuellement disponible et qu'il ouvre de nouvelles avenues explicatives, notamment dans le domaine des sciences humaines (Bunge, 2003, 2004 ; Sawyer, 2003).

Chapitre 1

La mise en place des fondements du néo-mécanisme

Il existe au moins deux consensus en philosophie des sciences : 1) il a existé un modèle standard de l'explication scientifique (le modèle D-N) et 2) tous s'entendent, à présent, sur sa faillite. Des paradigmes kuhnien à l'unificationnisme de Kitcher en passant par les théories sémantiques de Suppe ou la théorie pragmatique de Van Fraassen, le foisonnement des théories alternatives au modèle D-N sur la nature des sciences et la façon dont elles se développent a été impressionnant. Du propre aveu de l'un des premiers et principaux détracteurs de la philosophie de l'empirisme logique, W. Salmon, les courants actuels doivent être compris en réaction au modèle standard: « The 1948 Hempel-Oppenheim article marks the division between the prehistory and the history of modern discussions of scientific explanation. »⁹.

De façon générale, donc, le néo-mécanisme se caractérise par un tournant réaliste et naturaliste par rapport à l'approche linguistico-formelle des empiristes

⁹ Salmon (1989) *Four Decades of Scientific Explanation*, p. 10. Hempel et Oppenheim, selon Salmon, sont les premiers à s'intéresser explicitement au thème de l'explication scientifique et à en proposer un modèle formel, le modèle D-N, que nous allons examiner en détail.

logiques (nous nous pencherons essentiellement sur Hempel et évoquerons Reichenbach). Comme nous le constaterons, la plupart du temps, il est en complète opposition avec les thèses hempeliennes. L'empirisme logique est une approche linguistique (les théories sont des énoncés) et formaliste (les inférences sont des déductions ; d'où une tendance anhistorique de la justification des explications scientifiques (les facteurs psychologiques et le contexte de découverte sont exclus du cadre de la philosophie des sciences). Leur prétention est de trouver un modèle logique auquel se conforme toutes les explications causales scientifiques (sciences humaines incluses)¹⁰.

Le néo-mécanisme se caractérise par un tournant réaliste et naturaliste inspiré des philosophes Wesley Salmon, Herbert Simon et William Wimsatt : l'approche néo-mécaniste est donc en complète opposition avec les thèses hempeliennes. Particulièrement, la façon dont nous devons expliquer un évènement en sera profondément modifiée : expliquer n'est pas lier logiquement, par la déduction, deux énoncés mais trouver comment un arrangement matériel, à partir de la disposition de ses composants, peut occasionner un phénomène. Par exemple, pour Hempel, la formalisation de certaines régularités naturelles servent de lois sous lesquelles il suffit de subsumer des énoncés particuliers alors que pour les néo-mécanistes, les régularités sont le départ de l'investigation scientifique : quels sont les mécanismes sous-jacents qui permettent la régularité d'un comportement ? Examiner le pan informel de l'activité scientifique est, à n'en pas douter, une entreprise infiniment plus informative sur la science que le projet de l'empirisme logique. Le néo-mécanisme attire l'attention des philosophes des sciences sur le lieu même du travail effectif des acteurs scientifiques alors que l'empiriste logique se contente de reconstruire les raisonnements scientifiques de manière conforme au modèle D-N.

¹⁰ « Nous verrons que très souvent nos conclusions relatives aux méthodes et au raisonnement de la recherche scientifique s'appliquent aussi bien aux sciences de la nature qu'aux sciences sociales. Les mots « sciences » et « scientifiques » seront donc fréquemment utilisés pour désigner le domaine des sciences empiriques tout entier [...]. », in Hempel (1966) *Éléments d'épistémologie (Philosophy of Natural Science)*, p. 2.

Comme nous l'avons déjà évoqué, nous verrons qu'il existe deux grandes tendances parmi les théories néo-mécanistes. Malgré leurs différences, il faut souligner que le néo-mécanisme, tel qu'il a resurgit depuis une vingtaine d'années¹¹, est surtout occupé à comprendre l'aspect informel de la justification et, surtout, de la construction des théories scientifiques. Cela s'effectue par la mise en évidence des contingences empiriques qui produisent des événements, des types d'événements ou des processus (mais aussi par une perspective élargie sur la science, i.e. les contextes de production : histoire, institution, traditions de recherche, ...).

L'intérêt pour la construction des théories scientifiques, qui se manifeste par l'effort déployé pour comprendre la façon dont les chercheurs mènent leurs investigations, contraste grandement avec la « théorie créative » de l'empirisme logique. Le phénomène de la découverte a été conçu comme un phénomène mystérieux, le fruit du « génie » scientifique¹². Le thème de la découverte a été exclu du domaine de la philosophie des sciences suite à une distinction malheureuse entre contexte de découverte et contexte de justification. La conception logicienne de la philosophie de l'empirisme logique interdit l'investigation de domaines qui ne s'appréhendent pas aux moyens des outils d'analyse logique. Les outils d'analyse déployés par les néo-mécanistes résultent d'une approche pragmatiste et naturaliste des sciences et prennent la forme d'heuristiques de recherche et s'inspirent des méthodes expérimentales actuelles (chapitres 2 et 3). Contrairement aux empiristes logiques, à peu près tous les néo-mécanistes se sont penchés sur la construction de guide méthodologique pour la recherche, la plupart du temps à partir de cas tirés de

¹¹ Nous attribuons ce changement de perspective à W. Salmon (1984), qui l'attribue lui-même à Railton.

¹² Reichenbach (1951, p. 231) : « The act of discovery escapes logical analysis; there are no logical rules in terms of which a 'discovery machine' could be constructed that would take over the creative function of the genius. But it is not the *logician's* task to account for scientific discoveries; all he can do is to analyze relation between given facts and a theory presented to him with the claim that it explain these facts. In other words, logic is concerned with the context of justification. » (nous soulignons). Nous verrons, dans la section sur le tournant pragmatiste que Reichenbach a sous-estimé les 'discovery machine' de H. Simon. Popper (1959, p. 31) tient un discours identique, bien qu'il ne soit pas un empiriste logique.

l'histoire des sciences, leur but étant évidemment de faire partager les procédures des plus prolifiques scientifiques¹³.

Puisque notre intention dans ce chapitre est d'expliquer la réapparition du mécanisme dans le contexte de la philosophie de l'empirisme logique, nous montrerons comment s'est constitué un mouvement contre la philosophie des sciences de l'empirisme logique sur deux plans théoriques, celui du réalisme de W. Salmon et celui du naturalisme pragmatiste de H. Simon et W. Wimsatt¹⁴. Ces trois auteurs sont les instigateurs desquels se réclament tous les néo-mécanistes. Nous examinerons donc comment un tournant réaliste s'est accompli en réaction au modèle D-N, pour ensuite montrer comment, dans la mouvance de ce courant réaliste, s'est constituée une approche pragmatique et naturaliste de la philosophie des sciences. De plus, dans ce chapitre se profile la chute du physicalisme éliminativiste et son remplacement pour une ontologie physicaliste systémique non-éliminativiste. Plus précisément, les néo-mécanistes adoptent tous une théorie ontologique des systèmes complexes : la nature est organisée en niveaux compositionnels, c'est-à-dire une imbrication hiérarchique où chaque niveau (ou champs) entretient de fortes interactions internes mais aussi des relations, plus faibles, avec les niveaux supérieurs et inférieurs. La nature est donc organisée en niveaux qui possèdent chacun une

¹³ C'est ce qui explique la filiation du néo-mécanisme avec Simon et son programme du pragmatisme computationnel : « Simon's project is an ambitious research programme in artificial intelligence. He characterize the project, rather provocatively, as an attempt to construct a 'logic of scientific discovery'. The 'logic' that Simon seeks would be an explicit set of principles for reasoning and the conduct of inquiry which, when followed systematically, will result in the production of good scientific hypotheses and theories. To figure out how to produce good scientific theories, they study and try to simulate what good scientists do. » in Stich (1993) *Naturalizing Epistemology: Quine, Simon and the Prospects for Pragmatism*, p. 10-11.

¹⁴ Un autre courant d'opposition est incarné par T. S. Kuhn dont le principal mandat est de rectifier l'emphase mise sur l'importance des théories dans la dynamique des changements scientifiques en montrant comment une approche plus globale, qui inclue les facteurs socio-historiques (institution et paradigme de recherche), est requise. En marge de la figure médiatique de Kuhn se trouve des philosophes qui ont développés des notions comme les programmes de recherche (Lakatos), les champs (Darden et Maull, 1977 ; Darden, 1991), les pratiques (Kitcher, 1993) et les traditions (Laudan, 1977). Sans oublier le travail de Hacking (1983) qui tente de réhabiliter l'expérimentateur (en contraste avec le théoricien) comme figure essentielle, voir centrale, de l'avancée scientifique (dans la même veine, Latour et Woolgar, 1986, Latour 1987).

dynamique non réductible aux niveaux qui les composent. Nous verrons que le néo-mécanisme est particulièrement pertinent pour comprendre 1) le phénomène de l'émergence (chapitre 1) et 2) l'intégration de ces niveaux dans les explications scientifiques (chapitre 3).

I. Théorie hempelienne de l'explication

Globalement, selon Hempel, toute explication doit remplir deux exigences importantes (qu'il estime, évidemment, remplies par le modèle D-N) : l'exigence de pertinence et de testabilité¹⁵. La première exigence stipule que l'explication doit être informative relativement au phénomène à expliquer. La question substantielle qui sous-tend l'exigence de la pertinence est celle de la force explicative : quand reconnaissons-nous une explication et pour quelle raison ? Qu'est-ce qui permet une bonne explication ? La réponse de Hempel résidera, nous le verrons, dans la notion d'*expectabilité nomique*. La deuxième exigence, la testabilité, est la possibilité de concevoir, en principe, un test empirique de façon à infirmer ou corroborer l'hypothèse testée. Bien qu'importante, la condition de testabilité ne constitue qu'une obligation de contenu empirique pour toute théorie scientifique qui, comme nous le verrons, sont conçues d'abord comme des structures syntaxiques organisées en argument déductif. En fait, la condition de testabilité comme exigence de contenu empirique est due à la théorie sémantique vériconditionnelle qui devait prémunir les sciences des énoncés métaphysiques (ceux-ci n'ayant pas de conditions de vérité testables empiriquement, ils n'ont tout simplement pas de signification).

Le modèle déductif-nomologique se veut une structure linguistique formelle unique sur laquelle toute explication scientifique est construite¹⁶. Il est possible de distinguer deux éléments importants pour le modèle D-N : 1) une théorie linguistique

¹⁵ Hempel (1966) *Éléments d'épistémologie (Philosophy of Natural Science)*, p. 73-77.

¹⁶ Sur la portée du modèle D-N, rappelons Hempel (1966, p. 2) : « Nous verrons que très souvent nos conclusions relatives aux méthodes et au raisonnement de la recherche scientifique s'appliquent aussi bien aux sciences de la nature qu'aux sciences sociales. ».

(syntaxique) des théories et 2) une théorie déductiviste de l'inférence en science. Nous nous limiterons à une interprétation du modèle D-N¹⁷ basée sur ces deux théories à propos de l'explication scientifique, d'une part parce qu'elles font l'objet de critiques originales par les néo-mécanistes (chapitre 3), d'autre part, parce que les autres thèses ont été largement analysées et critiquées dans la littérature¹⁸ (la distinction entre langage observationnel et théorique, les principes internes et les principes de liaisons et la thèse de la Réduction Inter-Théorique (RIT ou IR pour Intertheoric Reduction¹⁹).

La théorie linguistique des théories pose qu'une explication est un argument, soit un ensemble organisé d'énoncés ou de propositions. Les entités constitutives d'une théorie scientifique ne sont que des énoncés²⁰. Les énoncés chargés de fournir l'explication sont appelés « explanans » et l'énoncé conclusif, ce qui doit être expliqué, est l'« explanandum » (E). La deuxième théorie est que l'argument liant l'explanans et l'explanandum est de type déductif. Pour ce faire, l'explanans est constitué d'au moins une loi universelle (T) et d'un énoncé singulier représentant des conditions initiales (C). La loi universelle est une conditionnelle universelle, c'est-à-dire qu'elle est de la forme « si x alors y », où « pour tout x , alors y » que l'on représente par l'expression $\forall x (Fx \rightarrow Gx)$. Expliquer revient donc à instancier un énoncé particulier sous une conditionnelle universelle et cette caractéristique formelle

¹⁷ Nous n'examinerons pas la variante Inductive-Statistique (I-S) du modèle D-N puisque considéré sous l'angle des deux thèses principales de l'explication des empiristes logiques, elle est similaire : le modèle I-S est un argument déductif très probable alors que le modèle D-N est un argument déductif certain (advenant de véritables lois de la nature). Une autre limite est celle de l'explication des lois dans le cadre du modèle D-N : nous nous attarderons uniquement aux explications concernant les événements singuliers, ce type d'explication étant favorisé par le néo-mécanisme.

¹⁸ Les critiques classiques du positivisme logique, présentées dans le collectif de Jacob (1996), qui s'appliquent à l'empirisme logique sont *Ce que les théories ne sont pas* (Putnam) et *Les deux dogmes de l'empirisme* (Quine). Pour une critique de la théorie de la réduction de l'empirisme logique, voir Churchland (1986), Craver (2001a), Bickle (2003), Machamer et Sullivan (2004), Schaffner (2004).

¹⁹ Cette expression est utilisée par Bickle (2003) ; Machamer et Sullivan (2004).

²⁰ Hempel ira même jusqu'à affirmer qu'un corpus entier de connaissances (scientifiques) peut être représenté comme un vaste ensemble K d'énoncés (1966, p. 70). Cette approche propositionnelle de la connaissance sera dénoncée par tous les néo-mécanistes (Glennan 1992 ; Machamer, Darden et Craver 2000 ; Bechtel et Abrahamsen 2005) qui adhèrent plutôt à une approche cognitive de la science (Giere, 2000). Nous aborderons ce sujet en détails avec Simon et Wimsatt.

motive la présentation du modèle D-N comme un modèle subsumptif²¹ – hérité de la physique. Le modèle D-N est schématiquement représenté ainsi :

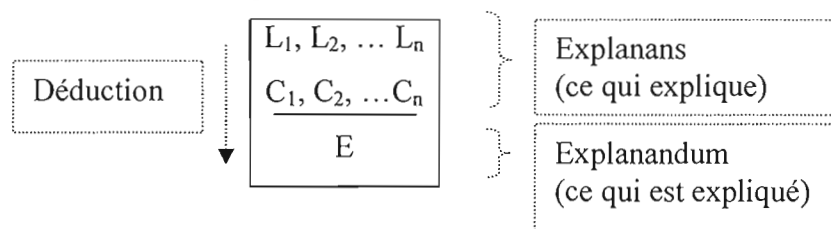


Fig. 1 : modèle D-N

Une remarque importante au sujet de la déduction est qu'il existe presque toujours des hypothèses auxiliaires qui permettent de déduire l'explanandum de l'explanans. Ainsi, sans l'arrière-plan fournit par ces hypothèses auxiliaires, presque aucune explication ne serait possible. L'existence de telles hypothèses pose le problème connu sous le nom de Duhem-Quine, bien que Hempel ne le reconnaisse pas sous cette appellation : quand les prédictions empiriques déduites de plusieurs hypothèses s'avèrent inexactes, il est difficile sinon impossible de savoir laquelle, ou lesquelles, sont fausses. En effet, si une seule hypothèse est testée et infirmée, il suffit simplement de la rejeter. Mais les hypothèses ne sont jamais testées seules : comment savoir, parmi la multitude d'hypothèses supposées (hypothèses principales et auxiliaires) laquelle est fausse, en presumant qu'une seule est fausse ? La discussion sur l'échec d'une expérimentation est une partie capitale du travail de scientifique qui est infiniment plus complexe que ne le laisse entrevoir le modèle naïf de la testabilité « une hypothèse – un test empirique »²².

Comme nous l'avons dit plus haut, une notion essentielle pour comprendre l'aspect formel du modèle D-N est l'expectabilité nomique. Comprendre, pour les empiristes logiques, c'est s'attendre à ce qu'un évènement ait lieu en vertu de son

²¹ Hempel (1966) lui-même le présente comme tel dans *Éléments d'épistémologie (Philosophy of Natural Science)*, p. 79.

²² Hempel ne s'attarde pas à réfléchir de façon plus systématique sur la façon dont les scientifiques s'y prennent pour résoudre ce problème (pour une approche computationnelle de résolution d'anomalies, voir Darden 1992, 1994, 1995 et 1998).

appartenance à une régularité naturelle. Toutefois, cette relation d'appartenance doit être comprise en terme de force logique entre deux énoncés :

A *DN* explanation answers the question “*Why* did the explanandum-phenomenon occur ?” by showing that the phenomenon resulted from certain circumstances, specified in $C_1, C_2, \dots C_k$, in accordance with the laws $L_1, L_2, \dots L_r$. By pointing this out, this argument shows that given the particular circumstances and the laws in question, the occurrence of the phenomenon *was to be expected*; and it is in this sense that the explanation enables us to *understand why* the phenomenon occurred.²³

L'explication d'un évènement particulier revient donc à identifier son appartenance à une catégorie générale d'évènement par une relation nomique. Un fil particulier de métal se dilate en fonction de l'augmentation de la chaleur parce qu'il est un métal et que tous les métaux ont cette caractéristique. Il est tentant de voir dans cette intuition une forme d'essentialisme, du moins une forme d'explication par appel aux propriétés naturelles intrinsèques à des genres naturels²⁴. L'idée fondamentale qui nous importe est que l'expectabilité nomique précise qu'expliquer n'est pas un rapport entre des entités du monde réel mais plutôt une implication logique (par la déduction) entre deux ensembles d'énoncés, l'explanans et l'explanandum.

L'expectabilité nomique se comprend encore mieux à la lumière de la théorie nomologique de la causalité²⁵ : « la théorie nomologique de la causalité réduit la

²³ Hempel (1965) *Aspects of scientific explanation and other essays in the philosophy of science*, p. 337, cité par Woodward (2003) *Making things Happen*, p. 153. Les italiques sont originaux.

²⁴ Cette impression peut être renforcée par la distinction loi universelle/généralisation accidentelle. Brièvement, les lois universelles supportent les conditionnelles contraires aux faits et les conditionnelles irréelles alors que les généralisations accidentelles ne le peuvent pas. Les lois sont de véritables régularités alors que les généralisations sont le fruit de découpages arbitraires. Malgré que les lois universelles sont fixées en rapport avec les théories établies par la communauté scientifique (Hempel, 1966, p. 89-90), le but de l'entreprise scientifique est de trouver de vraies lois : des lois de la nature. Par exemple, l'énoncé « toute sphère constituée d'or est inférieure à 100 000 kilogrammes » est vraie mais c'est une généralisation accidentelle, pas une loi de nature : rien dans la nature ne semble devoir interdire un telle possibilité. L'énoncé « aucune sphère constituée d'uranium est supérieure à 100 000 kilogrammes » n'est pas une régularité accidentelle en ce que les lois de la nature rendent impossible une telle chose (la masse critique de l'uranium est de quelques kilogrammes). Tiré de Salmon (1989) *Four Decades of Scientific Explanation*, p. 15.

²⁵ La théorie de causalité de Hempel est nommée théorie nomologique de la causalité par Barberousse, Kistler et Ludwig (2000) *La philosophie des sciences au XX^{ème} siècle*, p. 113.

relation causale à la relation entre une cause Fa et un effet Ga au fait que Fa est l'antécédent et Ga le conséquent de l'instance $Fa \rightarrow Ga$ d'une loi $\forall x (Fx \rightarrow Gx)$ qui lie la propriété F à la propriété G . »²⁶. Évidemment, la motivation principale pour limiter la causalité à une relation formelle entre deux énoncés ou variables est le parti anti-métaphysique hérité de la tradition empiriste (Hume). Comme nous l'avons annoncé dans l'introduction, nous verrons avec Salmon que le néo-mécanisme, en réclamant une théorie ontique de la causalité, réintroduit la métaphysique par le réalisme (la causalité est un certain type de processus du monde) dans la philosophie des sciences.

Une conséquence importante de l'aspect formel du modèle D-N est qu'il n'indique aucune contrainte temporelle de l'explanans sur l'explanandum (seulement des contraintes formelles). Par exemple, s'il s'agit d'expliquer la position de la Terre par rapport au Soleil durant la journée x à y heure en 1948, il existe deux méthodes. Dédire (E) à partir des données de 1947 ou dédire (E) à partir des coordonnées de la Terre en 1949 (les lois utilisées étant les lois du mouvement de Newton). Le second cas, où l'on déduit à partir d'un (C) postérieur, l'événement est quand même considéré comme une cause de la position de la Terre en 1948 : la théorie nomologique de la causalité reconnaît comme cause un énoncé particulier qui permet d'instancier une loi. Cette neutralité face à la directionnalité temporelle du modèle D-N a pour conséquence la fameuse thèse de la symétrie de l'explication et la prédiction : que l'on explique, que l'on prédise ou que l'on rétrodisse, on déduit (E) à partir de l'instanciation de (T) par (C) (ainsi que les hypothèses auxiliaires).

En résumé, le modèle D-N soutient que l'explication est de nature linguistique et formelle. Elle est linguistique car l'explication ne peut être qu'une relation entre

²⁶ *Idem*. Cette théorie de la causalité comme simple conjonction constante (logiciée) de deux événements est évidemment inspirée de Hume et comme ce dernier, Hempel exprime un refus de toute variante d'une connexion nécessaire (objective) entre les événements. Selon une expression, à la saveur Mackie, de Newton-Smith, « There is no causal glue cementing the events of the world together. », in Newton-Smith (2000) *A Companion to the Philosophy of Science*, p. 167. C'est précisément ce que niera Salmon.

des énoncés, sous la forme d'un argument et en langue naturelle ou en un quelconque langage artificiel (logique ou mathématique). Elle est formelle puisque expliquer est déduire l'explanandum de l'explanans.

II. Le tournant réaliste (Salmon)

Nous avons examiné comment Hempel pense pouvoir rendre compte de l'objectif principal de l'activité scientifique, l'explication, au moyen de l'analyse logique qui fournit des conditions formelles de validité. Salmon cherche à éclairer, lui aussi, ce qu'est une explication scientifique²⁷ mais il juge très sévèrement l'approche linguistico-formelle de son mentor :

It is my view that attempting to give a logical characterization of scientific explanation is a futile venture, and that little of significance can be said about scientific explanation in purely syntactical or semantical terms.²⁸

L'approche réaliste des explications scientifiques de Salmon est l'héritière du modèle D-N, bien qu'étant un successeur plutôt iconoclaste. De Hempel, Salmon conserve trois idées majeures : une attention accordée uniquement aux critères de justification de l'explication, l'idée sous-jacente selon laquelle la nature de l'explication scientifique est une problématique centrale de la philosophie des sciences²⁹ et finalement, l'idée que la physique est le modèle privilégié pour la philosophie des sciences. Plus précisément, pour les deux philosophes, les explications scientifiques sont de nature causale. L'apparent consensus ne peut être plus trompeur : Hempel soutient une théorie causale qui repose sur la nécessité logique de la déduction, alors que Salmon soutiendra une théorie réaliste, objectiviste

²⁷ « Our aim is to understand scientific understanding. We secure scientific understanding by providing scientific explanations; thus our main concern will be with the nature of scientific explanation. » in Salmon (1984) *Scientific Explanation and the Causal Structure of the World*, p. ix.

²⁸ *Ibid.*, p. 240.

²⁹ À ce titre, Salmon fournit une bibliographie chronologique (1948 à 1989) exhaustive sur la question de l'explication, dans Salmon (1989) *Four Decades of Scientific Explanation*, p. 196-219.

ou ontique de la causalité où la nécessité entre la cause et l'effet est une nécessité physique due à la structure causale du monde. Le conflit concerne la source de la nécessité que doit démontrer une explication : la nécessité selon Hempel est d'ordre logique (déduction dans un système logique) alors que pour Salmon, elle est due aux propriétés physiques de la matière (mise à jour de processus causaux dans un réseau causal). Débat fondamental, s'il en est un, puisqu'il concerne l'origine de la force explicative des explications scientifiques.

Ce tournant réaliste ou ontique de la causalité va radicalement transformer la théorie *causale* de l'explication : pour Hempel, ou pour les théories épistémiques³⁰ en général, expliquer est subsumer un énoncé particulier sous un énoncé universel alors que pour Salmon, ou pour les théories ontiques³¹ en général, et la philosophie causale-mécanique en particulier, expliquer un événement consiste à le situer dans un réseau causal (les causes étant étiologiques et constitutives, ou structurelles, nous y reviendrons dans la deuxième partie de ce chapitre).

Afin de comprendre ce tournant réaliste, nous examinerons 1) les objections de Salmon au modèle D-N et 2) la contribution positive de Salmon concernant le thème de l'explication scientifique, la philosophie mécaniste.

1) Le durcissement des positions

La version inférentielle de la conception épistémique, soit l'interprétation que fait Salmon de Hempel, se caractérise par le mode linguistique et formel d'appréhension des explications :

On this view, we can say that there is a relation of logical necessity between the laws and the initial conditions on the one hand, and the event-to-be-explained on the other – though it would be more accurate to say that the

³⁰ Épistémique parce que la force des explications réside entièrement dans une certaine forme d'opérations épistémiques (les déductions valides).

³¹ Ontique parce que la force des explications, qui sont néanmoins des opérations épistémiques, réside dans l'appel à des structures objectives du monde.

relation of logical necessity holds between the explanans-statements and the explanandum-statements.³²

Cette approche de l'explication, qui consiste à lier par la force logique de la déduction deux ensembles d'énoncés, va être critiquée de deux façons : une critique interne qui consiste à montrer que les explications ne sont pas nécessairement des arguments et une critique externe qui concerne la notion de causalité nomologique.

L'objection interne de Salmon, qui consiste à nier que toute explication est un argument, possède une visée générale contre la théorie de l'explication de Hempel : expliquer n'est pas une affaire de mise en rapport de phrases. L'objection aspire à miner l'aspect linguistique (syntaxique ou sémantique) de l'explication. L'argument se présente sous la forme d'une question dérangeante : dans un argument, pourquoi des énoncés non pertinents sont-ils fatals aux explications mais pas aux arguments ? Considérons l'argument suivant³³ :

- 1) Tous les hommes sont mortels
 - 2) Socrate est un homme
 - 3) Xanthippe est une femme
-

- 4) Socrate est mortel

Cet argument est parfaitement valide en dépit du fait que le troisième énoncé soit superflu. Poussé à l'absurde, l'on pourrait introduire une infinité de propositions sans aucun rapport avec le sujet sans pour autant affecter la validité de la déduction. Il en est tout autrement si l'on considère les énoncés 1)-4) comme une explication de la mortalité de Socrate. Mais tel quel, cet argument est une explication irrecevable de la mortalité de Socrate : le fait que Xanthippe soit une femme n'a rien avoir avec l'explanandum. Remarquons que le critère de pertinence de Hempel doit justement palier à ce problème des énoncés superflus, puisqu'ils sont fatals à l'explication. C'est

³² *Ibid.*, p. 16.

³³ Salmon (1989) *Four Decades of Scientific Explanation*, p. 102-103.

dire qu'il avait conscience de l'insuffisance des critères formels pour une théorie de l'explication causale mais, apparemment, il pensait se prémunir de cette insuffisance par son critère de pertinence³⁴. Souligner qu'expliquer quelque chose est fournir des facteurs pertinents à son apparition est sans doute vrai mais extrêmement peu informatif : si l'on pouvait savoir aussi facilement ce qui est pertinent de ce qui ne l'est pas, la recherche scientifique serait une chose bien facile. Il est clair que tout argument valide n'est pas une explication. Le point important le plus incriminant est, comme nous allons le voir à l'instant, que le modèle est fondamentalement déficient en ce qu'il ne permet pas de saisir ce qu'est une cause.

L'objection externe concerne un des défauts rédhibitoires du modèle : sa théorie de la causalité. Il semble qu'elle viole tout simplement la directionnalité temporelle intuitive à la notion de cause. Généralement, la cause d'un événement est considérée comme antérieure à celui-ci : il existe quelque chose qui appartient à la conformation de la cause et qui déclenche l'effet. Pas de causes antérieures, pas d'effets !

Toutefois, comme nous l'avons vu, les conditions formelles du modèle D-N n'interdisent pas que, factuellement, une cause (au sens nomologique) soit postérieure à l'effet (p. 21 : l'évaluation de la position de la Terre). Le contre-exemple³⁵ célèbre du drapeau de Bromberger montre comment la notion nomologique de cause peut être détournée tout en respectant les critères du modèle D-N. Imaginons que nous voulions mesurer l'ombre projetée par le mat d'un drapeau en fonction de la hauteur du Soleil. Un algorithme (T)³⁶ liant la hauteur du soleil (E) et la longueur de l'ombre (C) pourrait servir, en isolant la variable (E), à expliquer (E) en fonction de (C) : il

³⁴ Salmon a développé une théorie statistique de la pertinence, le modèle S-R. Celui-ci est le moyen épistémique dont nous disposons pour s'assurer d'avoir appréhendé les processus causaux qui influencent l'explanandum – ceux qui sont pertinents à l'explication. Nous ne l'aborderons pas puisque le point important pour nous est le tournant ontique, et non épistémique, qu'a permis Salmon.

³⁵ Pour un portrait complet d'autres contre-exemples voir Salmon (1989) *Four Decades of Scientific Explanation*, p. 46-50 ou, pour survol plus rapide, Humphreys (1989) *Scientific Explanation: The Causes, Some of the Causes, and Nothing But the Causes*, p. 300-301.

³⁶ (T) remplirait les conditions exigées par la définition de loi.

n'y aurait qu'à saturer (T) par l'argument « longueur de l'ombre » (C). Dans ce cas, la longueur de l'ombre du mat doit être considérée comme la cause (nomologique), grâce à (T), de la hauteur du soleil. En effet, rappelons-nous que la causalité nomologique est réduite à la relation entre une cause Fa et un effet Ga au fait que Fa est l'antécédent et Ga le conséquent de l'instance $Fa \rightarrow Ga$ d'une loi $\forall x (Fx \rightarrow Gx)$ qui lie la propriété F à la propriété G ³⁷. Tous s'entendraient à dire que le soleil cause l'ombre et non le contraire : la causalité est directionnelle.

Ce contre-exemple et de nombreux autres ont conduit Salmon (1984) à rejeter le modèle D-N comme étant une théorie inadéquate de l'explication scientifique causale. Devant ce fait, il a proposé de distinguer entre deux conceptions de l'explication scientifique, la conception épistémique et la conception ontique. Sa thèse principale est que le mode épistémique d'explication, particulièrement celui de Hempel qu'il nomme la version inférentielle³⁸, est déficient et que la conception ontique est une alternative plus viable³⁹. Ultimement, nous montrerons comment Salmon rejette le modèle D-N comme un modèle plutôt descriptif qu'explicatif, ce qui constitue une objection majeure à une supposée théorie de l'explication scientifique.

En résumé, l'approche causale-mécaniste est l'antithèse de l'approche épistémique sous deux aspects : 1) la nature de la nécessité de l'explication (épistémique ou ontique), 2) la façon d'expliquer un événement (globaliste ou particulariste). La théorie inférentielle de l'approche épistémique stipule que la nécessité procurée par l'explication est d'ordre logique plutôt que physique pour une approche ontique. Hempel pose qu'expliquer est toujours inclure logiquement un énoncé particulier sous un énoncé universel alors qu'une approche ontique tente plutôt de retracer les contingences matérielles qui ont produit l'évènement à expliquer : « A scientific explanation of some particular outcome will consist of

³⁷ Barberousse, Kistler et Ludwig (2000) *La philosophie des sciences aux XX^{ème} siècle*, p. 113.

³⁸ Les autres conceptions épistémiques sont la version informationnelle-théorique de Greeno et la version érotétique de van Fraassen.

³⁹ Salmon (1984) *Scientific Explanation and the Causal Structure of the World*, p. 239.

citing (some portion of) the causal process and interactions leading up to that outcome. »⁴⁰.

2) La philosophie causale-mécaniste

Salmon se définit comme un réaliste soutenant un physicalisme mécaniste, ce qui signifie que 1) nous avons un accès généralement fiable aux macrophénomènes dont l'existence est indépendante de l'esprit humain (réalisme), 2) cette fiabilité nous indique que le monde est composé d'évènement et de processus liés causalement (physicalisme) et que 3) ces liaisons causales forment des mécanismes (mécanisme). Plus précisément, Salmon est partisan d'une théorie causale objectiviste et particulariste de l'explication : « my account of scientific explanation is the emphasis I place upon spatiotemporally continuous causal connections [...]. »⁴¹. Expliquer revient à fournir les contingences du monde (sous la forme de mécanismes) qui sont causalement responsables de l'évènement. Il soutient aussi que l'ensemble des mécanismes peut être réduit à trois mécanismes fondamentaux, sur lesquels nous reviendrons. Une telle position impose un grand poids théorique sur la notion de cause et, plus largement, sur une théorie ontologique générale : qu'est-ce qu'un processus spatiotemporel continu de connexions ? comment peut-on l'identifier ? Pour Salmon donc, une théorie causale-mécaniste consiste essentiellement en la construction d'une ontologie des processus et des interactions entre les processus.

Cette perspective objectiviste ou ontique de la causalité soutient que la causalité est un phénomène physique dont l'existence n'est pas dépendante de son appréhension par l'esprit humain ou tout autre esprit. L'objectif de Salmon est donc d'offrir des hypothèses ontologiques sur ce qu'est une relation causale (ou connexion causale directe⁴²) mais aussi des hypothèses épistémiques sur la façon dont nous

⁴⁰ Woodward (1989) *The Causal Mechanical Model of Explanation*, p. 358.

⁴¹ Salmon (1984) *Scientific Explanation and the Causal Structure of the World*, p. 241-242.

⁴² Cette notion de causalité comme connexion directe (p. 261) sera reprise, quoique légèrement modifiée, par Glennan (chapitre 2) et critiquée par MDC (chapitre 3).

pouvons arriver à les saisir, autant pour la nécessité de la vie quotidienne (comme le reconnaît Hume) qu'afin de fournir des explications scientifiques causales-mécanistes. Or, la plupart des auteurs le rappelle systématiquement⁴³, toute approche ontique ou objectiviste de la causalité doit faire face au fantôme humien (Kitcher, 1989). Salmon situe d'ailleurs son projet philosophique à propos de la causalité comme une réponse à la position humienne concernant les relations causales⁴⁴.

Nous verrons, en premier lieu, l'objection de Hume à l'encontre de toute théorie causale objectiviste ; soit qu'un empiriste conséquent condamne la causalité comprise ontologiquement comme connexion nécessaire à un dogmatisme métaphysique. Deuxièmement, nous verrons comment Salmon pense avoir construit une théorie objectiviste non métaphysique (empirique) de la causalité. Nous finirons en exposant la théorie de l'explication qui découle de l'ontologie physicaliste et mécaniste.

Le fantôme écossais

Les positions sceptiques de la philosophie humienne ont jeté un important discrédit sur la notion de causalité dans une perspective empiriste. Pour citer Russell :

The law of causality, I believe, like much that passes muster among philosophers, is a relic of a bygone age, surviving, like the monarchy, only because it is erroneously supposed to do no harm. (1929, p. 180)⁴⁵

Hume est unanimement reconnu comme l'instigateur de la condamnation d'une compréhension empiriste et objectiviste de la causalité. Selon lui, la causalité est une relation psychologique liant deux impressions (simple ou complexes). Plus précisément, cette liaison mentale constitue une des trois lois de l'association des

⁴³ Salmon (1984, 1989) ; Woodward (1989) ; Humphreys (1989) ; Kitcher (1989) ; Glennan (1992, 1996, 2002) ; Kistler (1999) ; Barberousse, Kistler et Ludwig (2000) ; Newton-Smith (2000) ; Prinz (2001) ; Machamer, Darden, Carver (2000).

⁴⁴ Salmon (1984) *Scientific Explanation and the Causal Structure of the World*, p. 155.

⁴⁵ *Ibid.*, p. 137.

idées (impressions complexes) de l'entendement humain⁴⁶. Nous lions causalement des impressions ou des idées quand une impression ou idée est en conjonction constante avec une autre impression ou idée. C'est dire que la causalité est une opération épistémique a posteriori qui naît par une exposition redondante à des covariances, autrement dit, par une habitude ou une accoutumance. Ainsi, j'ai associé le pain avec le satiété puisque, habituellement, le pain me rassasie et je me représente, à tort selon Hume, que c'est le pain lui-même, de par ses caractéristiques physiques (ses « pouvoirs »), qui occasionne l'effet représenté par l'impression de satiété.

L'argument de Hume est d'ordre sceptique : nous n'avons accès qu'aux atomes sensoriels que sont les impressions fournies par nos sens. Or, l'expérience des liens causaux, les « connexions nécessaires », est une expérience inaccessible :

Il apparaît que, dans les cas isolés d'opération des corps, nous ne pouvons jamais, par l'examen le plus poussé, découvrir autre chose que la succession de deux événements, sans que nous soyons capables de comprendre la force ou le pouvoir qui fait agir la cause, ou la connexion qui la joint à son effet supposé.⁴⁷

Vouloir connaître les connexions causales est donc un projet métaphysique, c'est-à-dire qui ne peut être soumis au tribunal empiriste de notre connaissance.

L'intuition objectiviste opposée à cette approche épistémique de la causalité est bien illustrée par un exemple de Glennan : lorsque je démarre mon automobile, je ne crois pas qu'elle démarre parce que j'ai constaté une conjonction constante entre deux impressions (tourner la clef et le vrombissement du moteur) mais bien parce que je crois qu'il existe un mécanisme qui lie ma clef au démarreur qui fait lui-même partie du moteur. Mieux, elle est vérifiable (donc empirique) : « Furthermore, this [la connexion clef-démarrage] is not a secret connexion. I can look under the hood and see how the mechanism works. »⁴⁸.

⁴⁶ Les autres étant la ressemblance et la contiguïté dans le temps et dans l'espace, Hume (1748) *Enquête sur l'entendement humain*, p. 72.

⁴⁷ Hume (1748) *Enquête sur l'entendement humain*, p. 141.

⁴⁸ Glennan (1996) *Mechanisms and The Nature of Causation*, p. 50.

Malgré la sympathie que l'on peut éprouver pour la réponse de Glennan, cette façon de chasser le fantôme humien est inefficace. Reprenons le cas de l'automobile : si l'on affirme que l'on connaît le lien « clef-démarrage » en fournissant une analyse plus détaillée, comme « clef-démarreur-bougie d'allumage-ignition-démarrage », un sceptique s'empressera de demander de rendre compte non pas d'une connexion, mais de quatre. Bref, l'argument humien de l'inaccessibilité de la connexion nécessaire est récursivement inflationniste dans la mesure où la stratégie consiste à détailler une chaîne causale, en insérant des étapes intermédiaires, pour expliquer un phénomène. Cette stratégie n'a pour effet que de multiplier des relations à expliquer : « It seems that the Humean question never gets answered. »⁴⁹.

Ghostbuster

Salmon soutient que les conclusions négatives de Hume concernant une approche objective et empiriste de la causalité sont dues à l'ontologie phénoméniste et atomiste qu'il adopte. En effet, pour Hume, la causalité est une relation⁵⁰ entre deux événements mentaux indépendants (des impressions ou des idées). La théorie ontologique générale de Hume est donc que notre monde est une série d'impressions (simples ou complexes) liées de façon contingente entre elles par des lois d'association de notre esprit. La stratégie de Salmon est d'opter pour une théorie ontologique générale alternative : notre monde n'est pas une série d'événements connectés psychologiquement, mais une structure *sui generis* de processus qui interagissent de façon continue au plan spatiotemporel. La notion de causalité sera alors fractionnée en production et propagation (ou transmission) de processus causaux : à partir de trois types fondamentaux d'interaction entre les processus causaux, il est possible d'expliquer la production de nouveaux processus (causaux ou

⁴⁹ Salmon (1989) *Four Decades of Scientific Explanation*, p. 107.

⁵⁰ Hume affirme que la relation est d'ordre psychologique mais d'autres théorisations sont possibles. Par exemple, Mackie pense la relation en terme de conditions suffisantes (les conditions « INUS »).

non) et d'expliquer comment un processus est propagé à travers le temps et l'espace.

Un processus est décrit par ce que Russell nomme une chaîne causale :

A causal line may always be regarded as the persistence of something – a person, a table, a photon, or what not. Throughout a given causal line, there may be constancy of quality, constancy of structure, or a gradual change of either, but not sudden changes of any considerable magnitude. (1948, p. 459).⁵¹

Aucune limite précise de ce qui distingue un événement d'un processus n'est donnée, seulement que les événements sont assez localisés dans l'espace et le temps alors que les processus sont généralement plus étendus dans l'espace et le temps⁵². Salmon soutient que ce changement de considération à propos des entités fondamentales permet de comprendre comment la causalité, comme processus, répond au problème de la connexion entre les facteurs causaux et les effets résultants (le fantôme humien) : les connexions nécessaires sont connaissables et se sont des processus causaux⁵³.

En effet, il existe deux types de processus, les processus causaux et les pseudo-processus qui ne sont pas causaux. Plusieurs critères empiristes (qui sont postulés comme étant admissibles par tout empiriste) permettent d'identifier les véritables connexions causales du monde. Les processus causaux possèdent : 1) la capacité de transmettre de l'information et de l'énergie, 2) la capacité de transmettre une marque (une modification dans la structure) c'est-à-dire la capacité de transmettre ses propres caractéristiques (nous y reviendrons). Les pseudo-processus ne remplissent pas toutes ces conditions et jamais la deuxième, ce qui fait de cette dernière le critère décisif de distinction⁵⁴.

⁵¹ Salmon (1984) *Scientific Explanation and the Causal Structure of the World*, p. 140.

⁵² *Ibid.*, p. 139. Le débat concernant les éléments fondamentaux de l'ontologie (entité ou processus) sera abordé plus en détails au chapitre 3.

⁵³ « [...] causal processes constitute precisely the causal connections that Hume sought, but was unable to find. » in Salmon (1984) *Scientific Explanation and the Causal Structure of the World*, p. 147.

⁵⁴ Salmon a depuis (1997, 1998) changé de position en mettant l'accent sur le critère spécifique de l'énergie : un processus causal est celui qui transmet des quantités de grandeurs conservées (traduction de Kistler (1999), p. 75). Ce critère d'identification des processus causaux sera critiqué, comme nous le verrons au chapitre 3, par MDC (2000) pour son inapplicabilité en biologie/neurobiologie.

Salmon offre des exemples pour faire comprendre ce que sont un processus causal et un pseudo-processus mais celui de Woodward (2003) est plus éloquent :

Suppose that a cue ball, set in motion by the impact of a cue stick, strikes a stationary 8 ball with the result that the 8 ball is put in motion and the cue ball changes direction. The impact of the stick also transmits some blue chalk to the cue ball, which is then transferred to the 8 ball on impact. The cue stick, the cue ball and the 8 ball are causal processes, and the collision of the cue stick with the cue ball and the collision of the cue and 8 balls are causal interactions. Salmon's idea is that citing such facts about processes and interactions explains the motion of the balls after the collision; by contrast, if one of these balls casts a shadow that moves across the other, this will be causally and explanatorily irrelevant to its subsequent motion because the shadow is a pseudo-process.⁵⁵

La transmission de marque est illustrée dans l'exemple de Woodward par le talc bleu. Le talc présent sur la queue de billard est transmis à la boule blanche qui le transmet à la boule noire. Ce processus est causal car une fois initié, la trace de talc se transmet sans intervention externe, ce dont est incapable le pseudo-processus suivant : l'ombre de la queue de billard fusionne avec l'ombre de la boule blanche qui fusionne avec l'ombre de la boule noire. L'ombre de la queue de billard ne peut absolument pas transmettre le talc bleu ou tout autre marque quelconque, c'est un pseudo-processus.

Distinguer les véritables processus et connexions causales parmi les processus est donc l'objectif de Salmon puisqu'il opte pour une théorie causale de l'explication, où expliquer revient à identifier les causes de l'explanandum ; un critère qui permet de distinguer les processus causaux des processus non causaux et en fait un critère qui distingue les processus explicatifs des processus qui ne permettent pas l'explication. La démarcation ontologique entre processus causaux et pseudo-processus est le fondement de la distinction épistémologique entre les facteurs causaux explicatifs des facteurs causaux non explicatifs. C'est surtout le critère de la transmission de marque (le deuxième critère) qui est développé pour distinguer les processus causaux des pseudo-processus :

The ability to transmit a mark is the criterion that distinguishes causal processes from pseudo-processes, for if the modification represented by the

⁵⁵ Woodward (2003) *Making things Happen*, p. 351.

mark is propagated; the process is transmitting its own characteristics. Otherwise, the 'process' is not self-determined, and is not independent of what goes on elsewhere.⁵⁶

Ainsi, selon Salmon, les processus causaux sont les processus par lesquels l'influence causale (tant la production que la propagation) est diffusée de par le monde. Ces processus ne sont pas des forces occultes ou des reliques métaphysiques d'une époque révolue mais des entités repérables par leur capacité de transmettre leur structure de par elle-même, i.e., une fois produite, elle se transmet d'elle-même (comme une onde radiophonique à travers les airs). Cette caractéristique peut être nommée propriété d'autopropagation (PA) : elle indique que tout processus de cette nature possède, dans un certain contexte, la propriété de se propager de façon continue dans l'espace et le temps sans subir d'altération de sa structure – constituant ainsi un support pour l'information et une configuration stable pour la conservation de l'énergie. L'idée intuitive de Salmon est que tout processus qui persiste spatiotemporellement, sans intervention pour entretenir sa persistance, est un processus causal. Un processus qui ne persiste pas spatiotemporellement, sans intervention pour entretenir sa persistance, est un pseudo-processus, par exemple, l'ombre d'un objet physique en mouvement⁵⁷.

Toutefois, la récursivité de l'argument humien s'applique à la notion de transmission de marque : comment une marque est-elle transmise suite à l'effet d'une influence causale ? en quoi consiste la transmission d'une marque ?

Pour répondre à ces questions, Salmon élabore le principe de la transmission de marque (TM) en suivant l'idée fondamentale du processus, la propriété PA : il y a transmission de marque dans un processus lorsque en tous points entre deux points étudiés A et B, la marque est uniformément présente sans qu'il n'y ait eu d'intervention. Autrement dit, la transmission de marque a lieu lorsque la propriété

⁵⁶ Salmon (1984) *Scientific Explanation and the Causal Structure of the World*, p. 145.

⁵⁷ « [...] if we try to mark the shadow by modifying his shape at one point (e.g. by altering a light source or introducing a second occluding object), this modification will not persist unless we continually intervene to maintain it as the shadow occupies successive spatiotemporal positions. », in Woodward (2003) *Making things Happen*, p. 350.

PA est respectée de facto pour tout point considéré entre A et B. Le principe TM consiste en la réalisation de la propriété PA : un processus causal est tel qu'il peut transmettre sans dépendance par rapport à un autre processus sa propre structure à travers l'espace-temps. On peut comprendre la TM graphiquement :

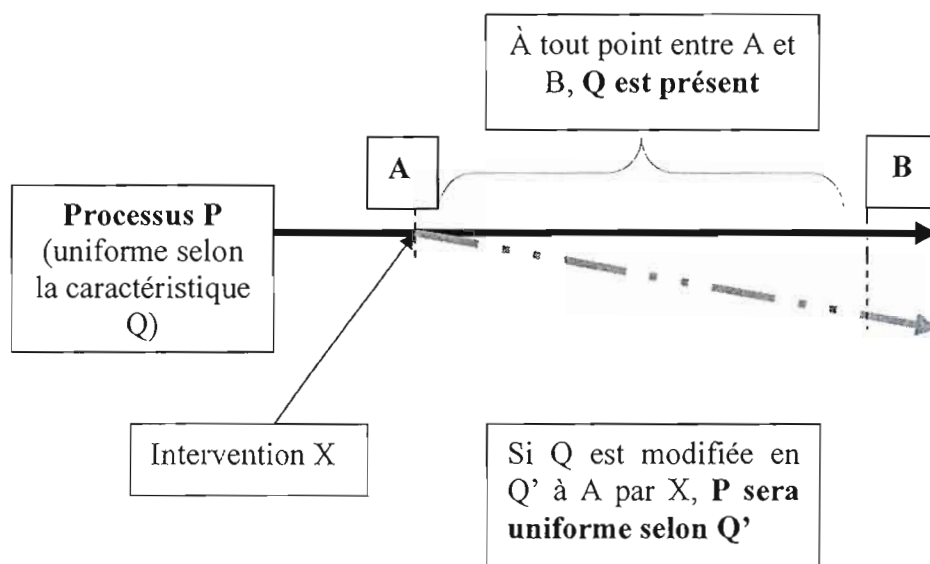


Fig. 2 : le principe de la transmission de marque

Cette modélisation de la TM est clairement contrefactuelle : 1) si l'intervention X a lieu, la caractéristique Q' sera uniforme dans P et 2) si l'intervention X n'a pas lieu, Q' ne sera pas uniforme dans P (ce sera la caractéristique Q). Salmon se distancie de toute approche pragmatique ou sémantique des contrefactuels et soutient que l'entreprise scientifique, par la méthode expérimentale, est un ensemble d'outils permettant de trouver si de tels processus opèrent dans le monde et si l'idée que nous nous en faisons est juste⁵⁸. Ce traitement

⁵⁸ Ce traitement empirique des contrefactuels a pour conséquence une conception intégrée des rapports entre sciences et philosophie : une fois nos affirmations formulées, il faut imaginer un test expérimental afin de les évaluer ou, de façon plus réaliste, reconnaître que l'évaluation des affirmations doit être complétée empiriquement dans le cadre des domaines scientifiques appropriés***. Woodward (2002) relève le défi d'analyser le mécanisme à l'aide d'une méthode

empirique des contrefactuels a pour conséquence une conception intégrée des rapports entre sciences et philosophie : une fois nos affirmations formulées, il faut imaginer un test expérimental afin de les évaluer ou, de façon plus réaliste, reconnaître que l'évaluation des affirmations doit être complétée empiriquement dans le cadre des domaines scientifiques appropriés⁵⁹.

Une dernière remarque à propos des processus causaux : la capacité de transmettre une marque (ou une structure) est l'indicateur que nous sommes en présence d'un processus qui peut propager une influence causale. Toutefois, la transmission de marques peut s'effectuer de multiples façons et c'est une erreur de conclure qu'un processus est non-causal s'il ne conserve pas un type particulier de marque. Frapper d'un bâton une tomate ou une balle en caoutchouc est sûrement un processus causal mais il n'est pas possible de les juger selon le même type de transmission. Une même énergie cinétique est transmise à la balle et à la tomate : la structure de la tomate sera désintégrée puisque elle est incapable d'absorber la force transmise par le bâton alors que la structure élastique de la balle le peut et va plutôt occasionner un transfert de la force cinétique – la balle sera projetée loin en avant. En un mot, l'influence causale se propage selon des facteurs antécédents (les processus causaux antérieurs) et constitutifs (les caractéristiques structurelles de l'entité qui est influencée par les processus causaux antérieurs). Toute évaluation de la TM doit donc prendre en compte ces facteurs (qui se traduisent ainsi par une explication étiologique et constitutive).

contrefactuelle inspirée de l'économétrie, incorporant ainsi dans la philosophie des sciences ce que Salmon avait tenté avec son modèle S-R.

***Salmon (1984, 1989) se livre souvent à l'exercice de penser des tests afin de résoudre des dilemmes théoriques et de souligner la nécessité de cette méthode pour la philosophie. En ce sens, il est un précurseur de la « philosophie expérimentale » dont se réclame un nombre grandissant de philosophes (Stich, Bechtel, Bickle,...). Le contraste avec la reconstruction rationnelle comme méthode philosophique s'avère extrême : pour les « expérimentaux », philosophie et science sont un continuum alors que pour les « reconstructeurs », le philosophe est celui qui valide les sciences en la restructurant selon les canons de la logique déductive.

3) L'explication causale-mécaniste

La remarque précédente conditionne la théorie mécaniste de l'explication de la philosophie causale-mécaniste. Toute explication est duelle et car elle comprend les aspects étiologiques et les aspects constitutifs qui mènent à la production de l'explanandum sous étude. Les aspects étiologiques sont les processus causaux antérieurs à la production de l'évènement *E* à expliquer et sans lesquels celui-ci n'aurait pu se produire. La figure ci-dessous expose graphiquement les explications constitutive et étiologique et comment celles-ci permettent de situer un phénomène à expliquer *E* dans un réseau spatio-temporel de processus.

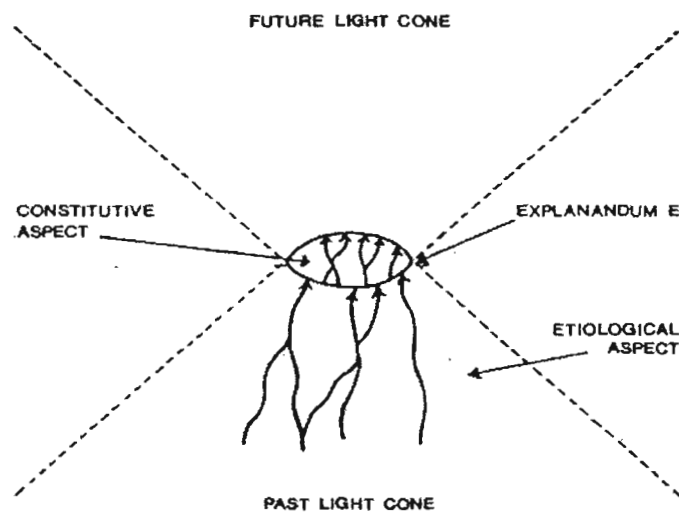


Fig. 3 : théorie mécaniste de l'explication⁶⁰

Toutefois, les explications étiologiques ne suffisent pas, comme nous l'avons montré dans l'exemple de la tomate et de la balle. Les idiosyncrasies de l'entité qui subit l'assaut des forces causales créent des réactions variables qui ne s'appréhendent que par une connaissance de la structure affectée. C'est le rôle des explications constitutives. Finalement, l'interaction dynamique entre les processus causaux et la structure particulière sera le lieu même où les mécanismes devront être élucidés.

⁶⁰ Salmon (1984) *Scientific Explanation and the Causal Structure of the World*, p. 275.

La conceptualisation des interactions entre les processus est donc la dernière notion clef de la philosophie causale mécaniste. Les interactions ont pour charge théorique la production de processus (alors que la propagation s'effectue par les processus). Les cas de figures sont limités à 1) les interactions de type X, 2) les interactions de type Y et 3) les interactions de type λ :

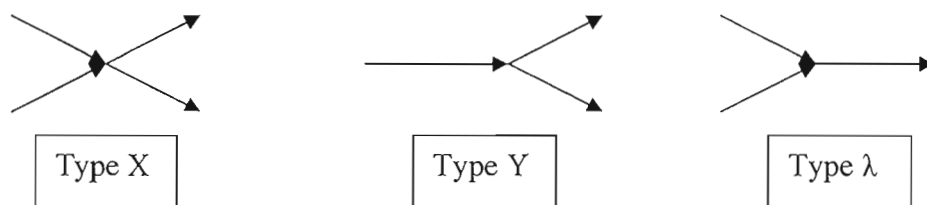


Fig. 4 : typologie des interactions fondamentales

Spécifions que les trois types présentés sont, en quelque sorte, des « briques de constructions » : elles peuvent être combinées pour rendre compte de phénomènes complexes. Ces types d'interactions permettent de construire des explications de la production de nouveaux processus à partir d'anciens processus : comment, pour reprendre l'exemple du démarrage d'une automobile, l'interaction entre la clé dirigée par l'intention d'un agent et le dispositif de démarrage peut produire le processus de démarrage (ce qui est un type λ : deux processus interagissent pour déclencher un nouveau processus).

Il existe deux genres d'interactions, celles qui sont causales et celles qui ne le sont pas (les pseudo-interactions). Les interactions causales respectent la TM : « When two processes intersect, and they undergo correlated modifications that persist after the intersection, I shall say that the intersection constitutes a *causal intersection*. »⁶¹. Ainsi, le type X représente deux processus P_1 et P_2 (causaux ou non) qui interagissent en produisant deux processus P_3 et P_4 . Les possibilités sont multiples. Premier cas de figure : P_1 et P_2 sont causaux et s'intermodifient de façon durable en

⁶¹ *Ibid.*, p. 170.

P_3 et P_4 (où P_3 et P_4 sont différents de P_1 et P_2). Deuxième cas de figure : P_1 et P_2 sont des pseudo-processus et ne s'intermodifient pas en P_3 et P_4 (donc P_3 et P_4 sont identiques à P_1 et P_2). D'autres cas sont possibles en vertu du fait qu'un processus causal ne modifie pas nécessairement un autre processus (causal ou non) : les processus causaux sont nécessaires mais non suffisants aux interactions causales.

Ces formes générales d'interaction sont les plus simples que l'on retrouve dans notre monde et chaque tentative de le comprendre peut se faire à partir de ces interactions fondamentales. Par exemple, un homicide à l'arme à feu (Salmon, 1984) : l'interaction (causale et en λ) entre une arme et une cartouche propulse une balle (processus causal) qui interagit avec le corps de la victime⁶². Bien évidemment, ce cas peut être examiné plus en profondeur sous différents aspects, comme le fait la balistique et la médecine légale. Mais d'un point de vue général et philosophique, la compréhension objectiviste de la causalité sert à appuyer l'hypothèse ontologique de la structure « en branche », formulée par Reichenbach, qui permet de comprendre comment un certain ordre (les types d'interaction) peut être créée par l'entremise de l'interaction entre processus (les structures arborescentes seront abordées dans la prochaine partie).

Une conséquence philosophique importante peut être déduite de cet ensemble de thèses posant que la structure du monde est due aux interactions combinées d'éléments (processus) fondamentaux. La philosophie mécaniste a pour conséquence une unification de nos théories scientifiques dans la mesure où tout phénomène dynamique est composé des mêmes éléments fondamentaux ; si le monde est construit à partir de trois types d'interactions, il faut nous attendre à voir disparaître la redondance ontologique au niveau explicatif. Donc, l'unification des sciences produite par le néo-mécanisme résulte d'abord d'un monisme ontologique : la sociologie, l'économie ou la biologie étudient des phénomènes physiques. Ensuite, le néo-mécanisme unifie l'épistémologie des sciences en plaidant pour une approche

⁶² *Ibid.*, p. 179.

causale des explications : la sociologie, l'économie ou la biologie étudient des phénomènes physiques par l'étude de leurs causes (physiques). Il ne s'agit donc pas de l'unification au sens des empiristes logiques où les sciences se réduisent à la physique.

Nos explications doivent reconstruire le réseau causal responsable de la production de l'explanandum, or, ce réseau causal est constitué de peu d'éléments fondamentaux (processus et interactions), donc nos explications sont constituées de ce nombre limité d'éléments fondamentaux bien que leurs combinaisons puissent être une source de grande difficulté de compréhension⁶³.

III. Le tournant naturaliste et pragmatiste (Simon et Wimsatt)

La compréhension de la résurgence actuelle du mécanisme s'envisage par rapport à une double réaction aux thèses de l'empirisme logique. Nous avons déjà montré comment la contribution de Salmon a permis, grâce à une théorie ontologique physicaliste centrée sur la notion de processus, un tournant réaliste. En un mot, la force explicative de nos théories scientifiques provient de la nécessité physique de la structure causale du monde plutôt qu'une nécessité logique dérivée de l'instanciation d'une loi de la nature. Cette considération ontologique par rapport à la source de la nécessité des explications scientifiques a pour conséquence principale que l'essentiel des facteurs explicatifs sont informels. Cette attitude réaliste est partagée par tous les néo-mécanistes et une majeure partie de leurs travaux consiste à éclairer les aspects informels des explications scientifiques. Ils ne nient pas qu'une approche formelle est

⁶³ En fait, nous verrons au chapitre 2 que Bechtel et Richardson cite les réseaux connexionnistes comme exemple de systèmes qui ne sont pas appréhendables par une approche mécaniste (bien que leur fonctionnement est supposé être mécanistique). Ces réseaux sont assez similaires à l'ontologie décrite ci-dessus puisque l'on saisit et maîtrise parfaitement une seule de ses unités mais lorsqu'un grand nombre d'entre elles sont combinées, les interactions deviennent trop complexes et deviennent « intractables ». Cette remarque impose une limite conceptuelle et opérationnelle/technologique sur le mécanisme : les entités ou processus doivent pouvoir être identifiés et suivis dans leur progression – ce qui exclut définitivement la microphysique du champs d'application du mécanisme.

nécessaire à une compréhension de l'entreprise scientifique, seulement que les aspects informels sont, de loin, plus informatifs. Les explications scientifiques doivent débusquer les processus causaux ainsi que leurs interactions pour pouvoir comprendre la manifestation des explanandum : la contrainte première des explications, c'est la structure du monde et non les carcans étroits de la logique déductive.

Un autre ensemble général de thèses est adopté par les néo-mécanistes, quelque fois nommé le naturalisme pragmatiste ou naturalisme faible (Stich 1993). Par naturalisme, il faut comprendre deux éléments : une attitude ontologique qui exclut toute entité surnaturelle (le matérialisme – émergentiste dans le cas de Simon, Wimsatt et les néo-mécanistes – est une position largement partagée) comme le phlogiston, l'âme ou les entités divines. L'histoire du monde naturel organique est conceptualisée par l'évolutionnisme qui ne recourt qu'à l'interaction d'éléments naturels (pas de causalité finale). L'autre élément est épistémologique : toute forme de connaissance a priori est refusée. Les connaissances humaines doivent provenir de l'interaction entre l'agent et le monde⁶⁴. En un mot, les naturalistes soutiennent que la réalité (physique et cognitive) s'épuise dans l'ordre causal du monde (Giere 2000b). Par pragmatisme, il faut comprendre le rejet d'une notion absolue de vérité : chacun poursuit ses recherches en fonction de valeurs. La cognition et le raisonnement sont des moyens par lesquels des objectifs peuvent être atteints. Une conséquence évidente et importante est que les stratégies de recherche peuvent différer selon les valeurs. Un pragmatiste s'intéresse essentiellement aux moyens permettant d'accomplir les objectifs fixés. Un projet pragmatiste se compose typiquement de quatre étapes⁶⁵ : 1) fixer des objectifs (en fonction des valeurs), 2) s'informer sur les méthodes les plus efficaces pour satisfaire les objectifs (identifier les acteurs les plus efficaces), 3) découvrir les stratégies de recherche et de raisonnement utilisés par ces acteurs et 4)

⁶⁴ Giere (2000b) *Naturalism*, p. 308. Cette façon de caractériser le naturalisme ouvre la voie aux approches développementales.

⁶⁵ Stich (1993) *Naturalizing Epistemology: Quine, Simon and the Prospects for Pragmatism*, p. 9-10.

tenter d'améliorer celles-ci. Tout travail scientifique est donc vu comme un processus cognitif, phénomène naturel au même titre que la fermentation de la levure : le naturalisme pragmatiste réhabilite la découverte scientifique comme un sujet digne d'investigation scientifique par l'intelligibilité qu'elle lui procure.

Parce que le projet du naturalisme pragmatiste s'intéresse à la compréhension des processus cognitifs responsables de la découverte en science, ce projet renverse l'approche de l'empirisme logique (rejet du contexte de découverte, concentration sur le contexte de justification) en intégrant la psychologie et l'intelligence artificielle dans l'épistémologie et la philosophie des sciences. Bien que les résultats de ce programme soient tangibles⁶⁶, nous nous préoccupons plutôt de présenter la conciliation du réalisme et du pragmatisme que représente H. Simon (et W. Wimsatt); toutes les pièces théoriques du puzzle néo-mécaniste seront mises en place pour exposer et problématiser les deux versions exhaustives du néo-mécanisme.

L'approche de Simon est, en principe, identique à Salmon : les agents scientifiques doivent se soumettre à ce qui semble être un ordre causal du monde. Nous avons vu comment Salmon conçoit l'activité scientifique : lier l'explanans et l'explanandum contrefactuellement au moyen des méthodes expérimentales de la science. Ce faisant, nous mettons au jour des mécanismes : un ensemble ramifié de processus causaux interagissant entre eux. L'activité scientifique a donc comme but l'établissement d'une explication causale-mécaniste, c'est-à-dire montrer comment la production d'un événement particulier dépend de l'interaction de facteurs antécédents et de facteurs constitutifs.

Cependant, deux différences importantes entre Simon/Wimsatt et Salmon doivent être soulignées et tiennent au fait que Simon et Wimsatt sont des pragmatistes en plus d'être réalistes. Ils déploient des efforts considérables pour comprendre et

⁶⁶*Ibid.*, p. 13 : « Simon and his coworkers have used laboratory studies of problem solving and 'rediscovery' in which talented students are asked to come up with a law or a theory that will capture a set of data [...], the protocols generated by the successful students can be used as another source of data against which simulation programs can be tested. [...] Simon and his group have produced impressive simulations of Kepler's discovery of his third law, Krebs's discovery of the urea cycle, and a variety of other important scientific discovery. ».

améliorer notre capacité à débusquer des structures causales, projet que Salmon, en tant que philosophe, laisse intégralement aux mains des scientifiques (et de la méthode contrefactuelle expérimentale). Salmon s'intéresse exclusivement aux types de relations objectives entre l'explanans et l'explanandum. À l'inverse, Simon s'est spécialisé dans la systématisation des procédures cognitives impliquées dans les découvertes scientifiques : c'est ce qu'il nomme science de la conception. Wimsatt (1980/1996), quant à lui, a initié les heuristiques de recherche que nous aborderons au prochain chapitre. La seconde différence est leur ontologie : la thèse de l'arborescence des systèmes complexes (où thèse de la quasi-décomposabilité). La nature est vue comme un ensemble de systèmes complexes organisés hiérarchiquement : les interactions intrasystémiques étant plus fortes que les interactions intersystémiques, l'étude isolée des systèmes est ontologiquement possible – ce qui, pour des êtres limités cognitivement, comme nous le verrons dans la prochaine partie, est une condition nécessaire d'intelligibilité.

1) La nouvelle ontologie

Historiquement, les objections au mécanisme ont été soutenues sur le terrain de la complexité et par rapport au monde organique et mental : la vie, et l'esprit, seraient trop complexes pour pouvoir être appréhendés à l'aide d'éléments aussi « simples » que ceux reconnus par les matérialistes. Le cas du dualisme de Descartes est le plus significatif : porte-étendard de la physiologie mécaniste, il ne pouvait concevoir que l'esprit soit une faculté explicable mécanistiquement à partir du cerveau, il fallait que l'esprit n'ait aucun rapport avec la matière (proposition on ne peut plus problématique), celui-ci est donc immatériel. Encore, la stratégie argumentative anti-mécaniste a très souvent pris la forme d'un appel à l'inalysable, à la primitivité indiscutable de la vie ou des facultés mentales. L'exemple des concepts de sensibilité et d'irritabilité de Bichat (qu'utilise aussi le célèbre vitaliste von Haller ainsi que l'École de Montpellier qui compte parmi ses rangs le célèbre

médecin Bordeu) est révélateur : la capacité des corps à réagir à leur environnement est une propriété fondamentale des organismes vivants au même titre que la gravitation pour tous les corps. La sensibilité et la gravitation sont inanalysables, elles sont des données « brutes ». Une autre stratégie anti-mécaniste plus actuelle est la doctrine de l'équipotentialisme⁶⁷.

La possibilité d'explorer des systèmes aussi complexes que les organismes est assez récente. Il n'y a qu'à consulter l'ouvrage classique *Introduction à la médecine expérimentale* de Claude Bernard (1865) pour s'apercevoir que la possibilité même d'une compréhension du vivant d'un point de vue matérialiste et mécaniste était si loin de faire l'unanimité qu'il nécessitait un vigoureux plaidoyer. L'explication donnée à cette opposition par Bernard concerne évidemment l'incroyable complexité des systèmes vivants :

Car, ainsi que je l'ai déjà dit ailleurs, la complexité due à l'existence d'un milieu organique intérieur est la seule raison des grandes difficultés que nous rencontrons dans la détermination expérimentale des phénomènes de la vie et dans l'application des moyens capables de les modifier.⁶⁸ [...] Le problème du médecin expérimentateur consiste donc à trouver le déterminisme simple d'un dérangement organique, c'est-à-dire saisir le phénomène initial qui amène tous les autres à sa suite par un déterminisme complexe, mais aussi nécessaire dans sa condition que l'a été le déterminisme initial.⁶⁹

Seule une conceptualisation puissante de la complexité et un perfectionnement technique et méthodologique pouvait permettre une résurgence du mécanisme.

La théorie de la complexité telle que présentée par Herbert Simon illustre de façon paradigmatique l'ontologie générale véhiculée par les sciences contemporaines. Le défi relevé par ce dernier est intimement lié à la quintessence de tout projet mécaniste : l'occurrence d'un phénomène doit être comprise comme le seul résultat

⁶⁷ Voir les pages 80 et 81.

⁶⁸ Bernard (1865) *Introduction à la médecine expérimentale*, p. 104.

⁶⁹ *Ibid.*, p. 136. Il est intéressant de noter la ressemblance entre cette caractérisation du travail de l'expérimentateur et la définition du mécanisme de MDC (chapitre 3) : « Mechanisms are entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions. » MDC (2000), *Thinking About Mechanisms*, p. 3. Cette idée que les mécanismes sont des ensembles structurés qui produisent un effet de façon constante est récurrente dans les descriptions de ce qu'est un mécanisme.

de l'organisation et la nature de ses composants⁷⁰ (ce que Simon nomme un réductionnisme de principe⁷¹). La propriété qui rend la contribution de Simon si précieuse est qu'il défend une thèse sur la complexité en tant que telle et non sur un système complexe particulier (organisme quelconque). Ses thèses sont donc de nature ontologique (et empirique) ; elles visent la mise au jour d'une architecture de la complexité :

Mon thème central est que la complexité prend fréquemment la forme d'une arborescence et que les systèmes arborescents ont quelques propriétés communes qui sont indépendantes de leur contenu spécifique. Je montrerai que l'arborescence est un des schémas structurels de base qu'utilise l'architecte de la complexité.⁷²

La ressemblance entre Simon et Salmon est, encore une fois, à souligner. Salmon, comme nous l'avons vu, a tenté de rendre compte objectivement de l'hypothèse ontologique de la structure « en branche » par l'entremise des concepts de processus et d'interaction. Il est possible de voir l'entreprise de Simon comme l'extension de l'hypothèse de la structure « en branche » : la structure en branche est arborescente. Nous allons examiner ce que signifie cette hypothèse de l'arborescence et montrer comment la propriété de la quasi-décomposabilité en résulte, propriété ontologique essentielle à la théorie mécaniste (nous verrons, au prochain chapitre, comment la nature ontologique des systèmes influence la recherche des mécanismes par l'intermédiaire des heuristiques).

L'idée maîtresse de Simon est que les systèmes complexes ont fréquemment une structure déterminée spécifique indépendamment de leur contenu : ils sont hiérarchiques ou arborescents⁷³. Cette affirmation signifie qu'un système complexe⁷⁴

⁷⁰ Simon pense possible d'étudier la complexité indépendamment de la nature des composants : l'organisation est pour lui l'élément informatif.

⁷¹ Simon (1996) *Les sciences de l'artificiel*, p. 303.

⁷² *Ibid.*, p. 321.

⁷³ Simon (1996) affirme que les systèmes physiques, biologiques, cognitifs, linguistiques et sociaux ont une même structure : ils sont arborescents.

est divisible en sous-systèmes où chaque sous-système peut être lui-même divisé, et ce, jusqu'à d'éventuels systèmes fondamentaux (composés d'entités fondamentales). La connaissance procède donc essentiellement de façon descendante (top-down). Par « hiérarchique », il ne faut pas comprendre qu'un système ordonne toutes ses parties et ses sous-systèmes : les interactions ne sont pas seulement descendantes mais aussi ascendantes. Autrement dit, il faut inclure dans le terme « hiérarchie » les systèmes complexes dont les interactions inter-niveaux sont bidirectionnelles (du système au sous-systèmes et des sous-systèmes au système). Deux remarques doivent être soulignées : 1) cette bidirectionnalité des interactions souligne clairement le caractère non-éliminativiste de cette théorie ontologique, ce qui permet 2) une continuité épistémologique entre les niveaux (certains facteurs neurobiologiques influencent la culture mais la culture peut aussi influencer les systèmes neurobiologiques).

Les niveaux d'analyse fondamentaux se réfèrent à des sous-systèmes considérés comme fondamentaux uniquement d'un point de vue pragmatique : les limites inférieures d'une science, ses sous-systèmes fondamentaux, peuvent faire l'objet d'une autre science à part entière ou une science peut elle-même reconsidérer la nature de ses sous-systèmes (et ses entités) fondamentaux⁷⁴. L'important est de garder à l'esprit que, malgré les découpes opérées par nos sciences, la réalité se présente plutôt comme un vaste ensemble de systèmes variablement intégrés. Autrement dit, les diverses analyses menées par les sous-divisions de l'entreprise scientifique ne doivent pas faire oublier que des limites épistémiques ne sont pas nécessairement des limites ontologiques, bien que Simon postule l'existence des niveaux ontologiques réels – comme nous allons le constater. Le « naturalisme pragmatiste » de Simon est clair : les limites disciplinaires sont souvent pragmatiques

⁷⁴ « [...] j'entends par système complexe un système fait d'un grand nombre de composants ayant beaucoup d'interactions. », *ibid.*, p. 320.

⁷⁵ Un exemple donné par Simon est la physique. La théorie de l'atome a connu de profonds changements au cours du XX^{ème} siècle, changements qui se sont effectués de façon descendantes : l'atome a d'abord été considéré dans son sens étymologique pour ensuite être divisé entre électron et protons, puis en électrons, protons et neutrons, ...

mais ce fait ne doit pas occulter la nécessité de rendre compte de la complexité du monde dans nos théories.

Les systèmes complexes sont donc hiérarchiquement organisés, c'est-à-dire que l'on peut les concevoir comme un encastrement dynamique de sous-systèmes. Toutefois, les interactions entre les éléments d'un même système sont souvent plus fortes et plus nombreuses que les interactions entre les éléments de systèmes distincts. Une telle propriété est due à plusieurs facteurs. Wimsatt, autre figure dominante de la caractérisation ontologique des systèmes complexes⁷⁶, nous offre un élément explicatif du fait que les interactions intrasystémiques sont souvent plus fortes que les interactions intersystémiques (ceci est valable pour les sous-systèmes) : la taille des entités est un facteur important dans la force des interactions entre systèmes. Des entités de taille similaire auront généralement des interactions plus fortes que des entités de tailles différentes. L'explication physique est que la taille d'un objet détermine un certain ratio entre la surface et le volume, ratio intervenant causalement dans la réactivité chimique des objets :

With the 'engineering paradigm' that we normally assemble complex systems out of simpler parts – a process that can be iterated – entities at successfully higher levels of organization tend to show roughly geometric increases in size.⁷⁷

Évidemment, la taille n'est qu'un indicateur, c'est-à-dire une condition parfois suffisante mais non nécessaire.

Simon et Wimsatt soutiennent donc que la nature est organisée en niveaux compositionnels, c'est-à-dire une imbrication hiérarchique où chaque niveau (ou champs) entretient de fortes interactions internes mais aussi des relations, plus faibles, avec les niveaux supérieurs et inférieurs⁷⁸. De tels systèmes sont dits quasi-

⁷⁶ Et ardent mécaniste : « [...] I believe that my surgery is ultimately more conservative, particularly in defending a liberalized (and **non-eliminative**) descendant of classical mechanistic materialism, and is also more in accord with actual scientific practice. », (nous soulignons), in Wimsatt (1994) *The Ontology of Complex Systems: Levels of Organization, Perspectives, and Causal Thickets*, p. 209.

⁷⁷ *Ibid.*, p. 233.

⁷⁸ Simon (1996) *Les sciences de l'artificiel*, p. 341.

décomposables : ils peuvent être étudiés partie par partie, suite à une décomposition et une localisation (nous exposerons les détails au chapitre 2). La quasi-décomposabilité est la conséquence de la force des relations intrasystémiques par rapport aux relations intersystémiques : il est possible de négliger, pour les besoins de l'étude, les relations externes d'un système pour pouvoir décrypter la structure interne d'une seule partie. Ce qui distingue quasi-décomposabilité et décomposabilité (tout court) est la présence ou l'absence d'interactions entre les composants. Si les composants d'un système n'interagissent pas entre eux, alors le système est décomposable. S'il existe des interactions faibles, il est quasi-décomposable et si les interactions sont fortes, il est minimalement décomposable. Nous verrons, au deuxième chapitre, comment Bechtel et Richardson ont élaboré la notion de décomposabilité minimale pour les systèmes dont les interactions sont fortes, le cerveau étant le paradigme d'un système intégré. Soulignons que la méthode mécaniste de décomposition et localisation des systèmes dépend étroitement du type de système étudié, particulièrement de la force des relations intrasystémiques (nulle, faible ou forte). Le plus souvent, la science s'occupe de systèmes complexes, fortement intégrés ou les interactions intrasystémiques sont faibles ou fortes. Nous reviendrons aussi sur les stratégies qui permettent au chercheur d'adapter son activité cognitive à la nature de son objet d'étude avec l'exposition des heuristiques de recherche. Pour le moment, il convient de remarquer qu'il sera possible d'étudier séparément un sous-système indépendamment du système hôte. Éventuellement, la connaissance de chaque partie étant avancée, il sera possible d'envisager les interactions intersystémiques afin de rendre compte, mécanistiquement, du comportement global d'un système complexe.

Être quasi-décomposable est une propriété importante pour nous, agents cognitifs devant naviguer dans un environnement parfois hostile et, plus philosophiquement, pour comprendre notre monde et nous-mêmes. En effet, la quasi-décomposabilité de certains systèmes complexes est une condition *sine qua non* de l'activité épistémique. Le raisonnement va comme suit. Nous sommes des êtres

grandement limités (mémoire, capacité d'attention, puissance de calcul, ...), ne pouvant circonscrire que des phénomènes d'une relative simplicité. La propriété de la quasi-décomposabilité nous autorise à étudier la réalité complexe d'une manière fractionnée parce que la nature elle-même est structurée en divers niveaux agrégés⁷⁹. Donc, la thèse ontologique des systèmes complexes hiérarchiques quasi-décomposables permet l'entreprise mécaniste (comprendre le complexe par l'entremise d'une interaction de mécanismes simples). Simon soutiendra même que les systèmes qui ne sont pas quasi-décomposables échappent à notre emprise cognitive dans le sens où de tels systèmes ne pourraient pas être saisis par nous : l'impossibilité de décomposer un système complexe nous obligerait à devoir maîtriser un nombre d'interactions largement supérieur à nos capacités mémorielles et computationnelles⁸⁰. Toutefois, il faut reconnaître qu'il est possible d'étendre nos limites mémorielles et computationnelles grâce à divers artefacts et modèles (téléscope/microscope, ordinateurs, ...). Mais il faut s'attendre à ce que cette externalisation de nos capacités soit limitée de la même façon que nos systèmes internes : si nous les avons construits, ils doivent obéir aux principes d'arborescence – sans quoi nous n'aurions pu les construire. À tout le moins, ces artefacts sont limités par l'exigence d'avoir une interface appréhendable pour nous (et donc de nos limites), sinon leur pertinence est nulle.

⁷⁹ Cette thèse ontologique a le mérite d'expliquer pourquoi différents domaines scientifiques peuvent étudier ces différents niveaux de façon relativement autonome (la sociologie peut se pratiquer indépendamment de la biologie, bien qu'il soit possible, et sûrement souhaitable, d'investiguer la sociologie d'une perspective biologique (comme le fait, très modestement, la neuroéthologie et certaines études biologisantes de l'épidémiologie de certains phénomènes sociaux).

⁸⁰ Simon (1996) *Les sciences de l'artificiel*, p. 357.

2) Les arguments en faveur de la nouvelle ontologie

En plus de naturaliser l'approche mécaniste en philosophie des sciences⁸¹, Simon et Wimsatt proposent deux arguments pour appuyer leur thèse ontologique, l'argument évolutionniste des horlogers (Simon) et l'argument perceptuel de la robustesse (Wimsatt).

Le premier argument en faveur de la thèse ontologique selon laquelle les systèmes complexes ont une structure arborescente est de type évolutionniste. Simon soutiendra qu'un système arborescent possède un avantage adaptatif tel qu'il supplanterait tout système compétitif non-arborescent – dans les limites temporelles imparties par l'histoire naturelle. Ainsi, il faudrait s'attendre à ce que les systèmes présents et soumis au mécanisme de la sélection naturelle soient de forme arborescente. L'argument est présenté métaphoriquement par deux horlogers Hora et Tempus. Tous deux étaient réputés pour leur admirable travail et les commandes affluaient. Toutefois Tempus fit faillite. L'explication de cette faillite réside en la méthode de construction des montres (d'environ 1000 pièces) utilisée. Hora les construit en assemblant dix sous-ensembles stables de cent pièces. Tempus, quant à lui, construit ses montres en une seule opération continue. Tempus et Hora étant des travailleurs autonomes, ils devaient répondre eux-mêmes au téléphone afin d'enregistrer les commandes. À chaque appel, Tempus doit interrompre la construction et perd le travail accompli (qui doit être continu) alors que Hora peut interrompre son travail sans perte de productivité car il monte des sous-ensembles plus petits et surtout, stables, exigeant une courte période de montage. Moralité : « le délai nécessaire pour l'évolution d'une forme complexe à partir d'éléments simples

⁸¹ Si toute créature cognitive doit fractionner la réalité pour la comprendre, la méthode de décomposition mécaniste est une extension épistémologique (et méthodologique) analogue au processus naturel de la cognition. Cette continuité entre la cognition commune et l'activité cognitive des scientifiques est revendiquée par Simon et Wimsatt, et tout naturaliste, par le biais de la thèse selon laquelle 1) le raisonnement est processus naturel et 2) la science est une activité de résolution de problèmes.

dépend de façon critique du nombre et de la distribution des formes stables intermédiaires potentiellement disponibles. »⁸².

Un critique soucieux de réalisme biologique pourrait remarquer que l'évolution ne procède pas toujours par assemblage de sous-ensembles dans un ensemble⁸³ mais aussi par multiplication et spécialisation de sous-systèmes au sein d'un système (comme l'effectue l'ontogenèse). L'évolution peut procéder par assimilation-intégration, comme dans le cas de la théorie de l'endosymbiose : les cellules eucaryotes et procaryotes ont assimilé et intégré nombres de leurs organites (mitochondries). Une deuxième stratégie est la multiplication-spécialisation : l'œil, par exemple, aurait débuté par de simples cellules photosensibles pour ensuite devenir le système visuel complexe actuel. Cette évolution aurait été possible par une multiplication de modules spécialisés (le cristallin et les muscles pour pouvoir obtenir un focus, une aire corticale spécialisée, ...). Simon remarquera que cette deuxième « stratégie » de l'évolution possède aussi un caractère arborescent. Par exemple, le système digestif qu'on divise habituellement en bouche, œsophage, estomac et intestins. Toutefois, à l'encontre de Simon, il faut reconnaître que les deux stratégies exemplifiées sont complémentaires et existent aussi de façon indépendante. Seulement, la stratégie d'assimilation-intégration est plus propre au monde bactériologique où les bactéries, par exemple, s'échangent du matériel génétique alors que la stratégie multiplication-spécialisation est plus propre aux organismes multicellulaires.

Dans tous les cas, l'argument de Simon est que les systèmes arborescents, indépendamment de leur mode d'évolution, ont un potentiel d'évolution considérablement plus important que les systèmes non arborescents (comme la montre de Tempus). Ce fait s'explique par l'optimisation locale que la quasi-

⁸² Simon (1996) *Les sciences de l'artificiel*, p. 332.

⁸³ Bien que des théories de l'évolution appuient cette possibilité : notamment la théorie l'endosymbiose de L. Margulis (1970) qui postule que les cellules ancestrales à l'origine des eucaryotes et des procaryotes ont été obtenus par un genre de phagocytose (des cellules originelles, micelles, aurait intégré des organites existant indépendamment, comme les mitochondries – qui ont leur propre ADN).

décomposabilité permet : chaque niveau, étant relativement autonome, peut devenir plus efficace sans nécessiter un réarrangement global – ce type de réarrangement étant généralement voué à la non-viabilité. Ainsi, tous les niveaux soumis à la pression évolutive peuvent se maximiser de façon quasi-indépendante. Bref, la structure arborescente de nombres de systèmes complexes est rendue compréhensible par sa capacité d'évolution rapide comparativement à un système de type « Tempus ».

En résumé, l'hypothèse de la quasi-décomposabilité est donc d'abord considérée comme une affirmation empirique généralement vraie (les systèmes complexes sont des arborescents et ont la propriété de quasi-décomposabilité) dont la présence peut être expliquée de façon évolutive par le haut degré d'adaptabilité qu'il possède.

Il existe un autre argument de Wimsatt en faveur de l'existence des niveaux d'organisation hiérarchiques et du fait qu'ils sont quasi-décomposables. Nous avons un accès généralement fiable aux mésophénomènes dont l'existence est indépendante de l'esprit humain⁸⁴. Wimsatt présente un portrait général magistral de l'ontologie telle que l'envisage tous les néo-mécanistes :

Ontologically, one could take the primary working matter of the world to be causal relationship, which are connected to one another in a variety of ways – and together make up patterns of causal networks. These networks should be viewed as a sort of bulk causal matter – an undifferentiated tissue of causal structures – in effect, the biochemical pathways of the world, whose topology, under some global constraints, yields interesting forms.⁸⁵

Wimsatt opte donc pour un appareil ontologique constitué de niveaux compositionnels, c'est-à-dire des divisions hiérarchiques de choses organisées par des

⁸⁴ « In a way, then, this analysis has something in common with folk psychology and some of the basic assumptions of ordinary language philosophy : like them, it takes for granted that the world we see, live in, respond to, and act upon, is too important, too central to our way of being, to be dismissed. » in Wimsatt (1994) *The Ontology of Complex Systems: Levels of Organization, Perspectives, and Causal Thickets*, p. 219.

⁸⁵ *Ibid.*, p. 220.

relations méréologiques⁸⁶. Les niveaux sont des « strates » de fortes interactions entre les propriétés des entités qui les constituent : ce fait s'explique, nous le rappelons, en partie par le facteur de la taille.

L'argument qui soutient que la nature est distribuée en niveaux relativement distincts et intégrés repose sur un critère ontologique général d'ordre perceptuel : il faut considérer comme réel ce que dénotent nos perceptions robustes⁸⁷, or, la stabilité des niveaux compositionnels constitue une référence exemplaire de robustesse, donc, il faut considérer comme réels les niveaux compositionnels. En effet, pour Wimsatt, il est essentiel de reconnaître que notre expérience humaine du monde est centrale à notre conception du monde et que nous établissons notre connaissance scientifique (et notre connaissance générale) de façon conforme au modèle des niveaux d'organisation de type arborescent, c'est-à-dire que les explications s'effectuent la plupart du temps dans la perspective d'un niveau particulier d'organisation. Une façon naturelle de comprendre ce fait à propos des explications est de reconnaître que les niveaux sont des strates ontologiquement stables et régulières qui autorisent à nos systèmes perceptifs l'opportunité de construire des perceptions robustes.

Cet argument de Wimsatt est typiquement mécaniste en ce qu'il s'appuie sur notre expérience mésophénoménale pour former, par analogie, un modèle général de compréhension de la nature qui pourra être étendu aux sphères mondaines (comme il le fut d'abord pour les phénomènes célestes). Le modèle mécaniste possède un double avantage : il est simple et fécond parce qu'il est supposé être fondé sur une structure du monde⁸⁸ et il est une norme d'intelligibilité de la nature. La notion de robustesse est donc centrale : c'est une propriété censée nous indiquer quelles sont les

⁸⁶ *Ibid.*, p. 222.

⁸⁷ Est robuste toute perception qui s'avère constante malgré la diversité des voies d'accès empruntées : une perception robuste sera trans-individuelle et trans-modale (détectée par plusieurs sens et même plusieurs artefacts technologiques).

⁸⁸ Wagner, dir. (2002) *Les philosophes et la science*, p. 131. Simple car tout phénomène est en principe explicable en faisant appel aux entités (et leurs propriétés) qui les constituent et à la façon dont elles interagissent. Fécond car le mécanisme est une doctrine ouverte aux avancées scientifiques et que, historiquement, chaque version du mécanisme a contribué grandement à l'avancée des connaissances.

perceptions qui reflètent une entité stable du monde. Autrement dit, la robustesse est un indicateur fiable (et non certain⁸⁹) qu'une perception robuste d'un objet est la conséquence d'une interaction réelle entre l'objet externe et l'organisme percevant. Bref, une perception robuste est l'effet d'une cause externe robuste : donc une cause « réelle ». Toute perception qui nous indique que l'on peut mesurer, définir et même produire un objet sert à renforcer notre croyance en la réalité de l'objet. Plus un objet est robuste, plus nous croyons avoir affaire à un authentique objet, au sens réaliste du terme⁹⁰.

La force persuasive de cet argument se veut double. D'abord, les informations apportées par les systèmes sensoriels sont les seules auxquelles nous avons accès en tant qu'organisme (ce point de vue est la composante épistémologique du naturalisme, un empirisme fort). Refuser cette unique source de notre connaissance du monde extérieur nous conduit assurément à un scepticisme extrémiste et stérile. Ensuite, les pressions démesurément longues de l'histoire naturelle sont garantes d'une relative fiabilité des systèmes sensoriels et ce, pour des raisons évolutionnistes évidentes⁹¹. L'étonnante efficacité pratique de certains organismes à détecter, à se coordonner dans le temps et l'espace et, finalement, à survivre, peut être interprétée comme un gage d'une maîtrise étendue de la détection et de la manipulation de

⁸⁹ La certitude est une exigence épistémique inatteignable pour Wimsatt.

⁹⁰ « We feel more confident of objects, properties, relationships, etc. which we can detect, derive, measure, or observe in a variety of independent ways because the chance we could be simultaneously wrong in each of these ways declines with the number of independent checks we have. », Wimsatt (1994) *The Ontology of Complex Systems: Levels of Organization, Perspectives, and Causal Thickets*, p. 213.

⁹¹ Wimsatt propose un argument établi suite à une recherche sur des relations entre l'utilisation des informations de type visuel et tactile : le cerveau semble généralement privilégier les informations visuelles, toutefois, si des informations tactiles invalident les informations visuelles (ce que l'on touche ne s'accorde pas à ce que l'on voit), les informations tactiles seront privilégiées. Ceci s'explique fonctionnellement par le danger que représentent les illusions systématiques inhérentes au système visuel. L'existence d'un tel mécanisme de consolidation de l'information par référence croisée indique vraisemblablement que le fonctionnement global de l'organisme est orienté vers l'optimisation de la robustesse des informations perceptuelles. (la recherche en question est de I. Rock et C. S. Harris, *Vision and Touch*, Scientific American 216 (1967) 96-104 cité dans Wimsatt (1994) *The Ontology of Complex Systems: Levels of Organization, Perspectives, and Causal Thickets*, p. 215).

particularités importantes du monde⁹². Bref, nos systèmes perceptuels constituent la source première et relativement fiable de notre connaissance du monde, spécifiquement les perceptions robustes.

En résumé, Wimsatt défend l'idée que notre accès aux mésophénomènes par les perceptions robustes est fiable. Or, nous percevons le monde comme étant constitué de niveaux compositionnels, c'est-à-dire des divisions hiérarchiques de choses organisées par des relations méréologiques. Cette perception de la nature est robuste : la nature, selon toute vraisemblance, doit être comme nous la percevons⁹³.

Ce premier chapitre visait à montrer comment s'est constitué le néo-mécanisme comme thèse ontologique et épistémologique par rapport aux théories de l'empirisme logique. Nous avons constaté que de l'approche linguistico-formelle inspirée de la physique a été contestée sous deux aspects fondamentaux, celle de la source explicative des explications et sur l'importance des facteurs pragmatiques dans la recherche et la découverte en science. Le premier aspect réside en la contribution réaliste de Salmon. En un mot, la force explicative de nos théories scientifiques provient de la nécessité physique de la structure causale du monde plutôt qu'une nécessité logique dérivée de l'instanciation d'une loi de la nature. Cette considération ontologique par rapport à la source de la nécessité des explications scientifiques a pour conséquence principale que l'essentiel des facteurs explicatifs est informel. Cette attitude réaliste est partagée par tous les néo-mécanistes et une majeure partie de leurs travaux consiste à mettre à jour les aspects informels des explications scientifiques. Le second aspect, le naturalisme pragmatiste, implique (notamment) une réhabilitation du phénomène de la découverte scientifique comme un sujet digne d'investigation scientifique. Tout travail scientifique est donc vu comme un processus

⁹² Pensons aux espèces migratoires.

⁹³ Hume a déjà proposé un argument similaire pour expliquer l'impression de covariation stable entre des objets et nos perceptions : si nos perceptions affichent une certaine constance, c'est que le monde lui-même possède une certaine régularité. L'on reconnaît l'argument réaliste « no miracle ». Il est intéressant de remarquer que même le phénoménisme radical de Hume s'appuie sur un réalisme externe.

cognitif de résolution de problèmes. Parce que le projet du naturalisme pragmatiste s'intéresse à la compréhension des processus cognitifs responsables de la découverte en science, ce projet renverse l'approche de l'empirisme logique (rejet du contexte de découverte, concentration sur le contexte de justification) en intégrant la psychologie et l'intelligence artificielle dans l'épistémologie et la philosophie des sciences. Mais le naturalisme pragmatiste inaugure surtout une nouvelle ontologie non-éliminativiste qui sied infiniment mieux à la conjoncture actuelle des sciences : les sciences biologiques ayant acquis une importante légitimité scientifique, il a fallu intégrer les entités et propriétés biologiques dans notre ontologie scientifique – intégration effectuée par l'entremise de la théorie des systèmes complexes et hiérarchiques.

Les similarités relevées entre les brefs aperçus historiques sur le mécanisme de Descartes et de Bernard et le mécanisme récent de W. Salmon, H. Simon et W. Wimsatt nous permettent d'extraire une stratégie générale importante du mécanisme. Tous croient que le mécanisme (et le néo-mécanisme) en tant que théorie de l'explication et méthode d'investigation scientifique est légitimé par une théorie ontologique. Si nous décomposons un phénomène pour comprendre comment il est produit à partir de l'organisation et la nature de ses composants, c'est que la structure de la réalité est ainsi faite : elle est constituée de multiples niveaux organisés compositionnellement imbriqués, structure qui rend possible la quasi-décomposabilité. L'emphasis sur la valeur de cet argument varie au sein de la communauté des mécanistes. Certains préfèrent jouer de prudence en n'affirmant pas une thèse ontologique aussi forte et mettent plutôt l'emphasis sur l'aspect heuristique du mécanisme, soit l'explication de mécanismes complexes par l'interaction de mécanismes simples. En résumé, la stratégie mécaniste est, le plus souvent, un appel à se conformer à son objet de recherche : le monde est quasi-décomposable, décomposons-le.

Chapitre 2

Première théorie exhaustive du néo-mécanisme (Glennan/Bechtel/Richardson)

Nous avons montré comment les empiristes logiques ont défendu une approche formelle des explications scientifiques en les réduisant à des instanciations de lois universelles. Le travail philosophique consiste donc, pour eux, à axiomatiser les théories scientifiques afin de mettre en évidence leur structure logique. Pour cette raison, ils considèrent le travail du philosophe comme étant externe et normatif : la genèse et les méthodes du scientifique sont légitimes dans la mesure où les résultats produits respectent la logique déductive du modèle D-N. Pour leur part, les néo-mécanistes s'entendent pour affirmer que l'essentiel du travail scientifique véhicule ses propres règles de légitimité et s'effectue principalement sur un plan informel : le néo-mécanisme développe les patterns mécanistes informels qui lient l'explanandum et l'explanans⁹⁴. Ainsi, le néo-mécanisme repose davantage sur une approche interne et très descriptive de la science sans refuser, pour autant, une contribution normative. Les néo-mécanistes imposent certains critères normatifs 1) par dérivation de la

⁹⁴ Craver (2002) *Structure of Scientific Theories*, p. 73, in Machamer et Silberstein (2002) *Philosophy of Science*.

pratique normative scientifique elle-même et 2) par l'imposition de critères ontologiques à la notion de mécanisme (Glennan/MDC).

Après avoir montré, au chapitre précédent, comment cette entreprise est sympathique au réalisme et au naturalisme, il nous faut à présent illustrer comment s'articule cette approche mécaniste des explications scientifiques. Ce chapitre constitue deux tentatives de réponses à plusieurs thèmes centraux de la philosophie des sciences qui sont applicables à une partie importante des disciplines scientifiques : 1) comment comprendre le concept même de mécanisme, 2) comment se présente une telle explication, 3) comment l'évalue t-on et 4) comment s'élabore la recherche scientifique pour parvenir à découvrir des explications (mécanistes) ?

Chacune de ces deux tentatives de réponse représente une extension du programme mécaniste : sous sa dimension ontologique et épistémique (Glennan) ainsi que sous sa dimension méthodologique et heuristique (Bechtel et Richardson).

La première extension touche à la définition même de la notion de mécanisme et d'explication mécaniste. La tâche principale de son travail précurseur est de remplacer la métaphore de la machine⁹⁵ par une caractérisation explicite du concept de mécanisme⁹⁶. Très carnapien dans l'esprit, le projet de Glennan, développé inauguralement dans *Mechanisms, Models, and Causation* (1992), indique une faiblesse assez problématique du projet de Bechtel et Richardson – celle d'avoir développé leur heuristique de recherche sur une notion intuitive et mal définie de mécanisme. Que cette omission est problématique est clair : nous avons vu, dans le premier chapitre, que le mécanisme soulève des enjeux ontologiques et épistémologiques profonds – les ignorer est simplement impossible ou alors, superficiel⁹⁷. Son projet s'articule sur deux modes d'analyse, ses positions

⁹⁵ *Ibid.*, p. 4 : « Mechanism involves commitment to the principle that the world is like a machine and that natural phenomena should be explained in something like the way in which behavior of machines is explained. ».

⁹⁶ *Ibid.*, p. 22.

⁹⁷ Bechtel (2005a et 2005b) a d'ailleurs consacré, ultérieurement, quelques articles à définir la notion de mécanisme et l'explication mécaniste ; toutefois, sa proximité avec MDC est si grande que l'aborder ennuerait le lecteur. L'originalité de Bechtel, par rapport aux autres mécanistes, est qu'il

ontologiques sur le concept de mécanisme et ses positions épistémologiques sur l'explication mécaniste. Le projet ontologique vise à préciser quelles entités peuvent être légitimes comme partie (ou composante) d'un mécanisme et son projet épistémologique vise à défendre que les explications mécanistes, via des modèles mécanistes, sont des explications en vertu du fait qu'elles sont « similaires » à la réalité – nous présenterons donc sa théorie de la similarité.

La seconde extension, opérée par Bechtel et Richardson, met l'emphasis sur l'aspect heuristique du néo-mécanisme dans le contexte de la découverte scientifique : l'investigation des mécanismes et la construction des explications mécanistes se fait par l'intermédiaire de méthodes de décomposition des systèmes complexes et de la localisation de ses composants. Bechtel et Richardson tenteront d'extraire des guides de recherche adaptés au type de système étudié. Leur travail est donc en continuité avec les hypothèses de Simon et Wimsatt.

Bien que les clarifications de Glennan soulignent des faiblesses de celui de Bechtel et Richardson, nous jugeons plus judicieux de considérer leurs analyses de façon complémentaire car, réunies, elles couvrent le contexte de découverte des mécanismes et offrent une description, accompagnée de plusieurs contraintes, de l'ontologie qu'elle suppose et de l'épistémologie qu'elle applique. Deux arguments peuvent être invoqués pour justifier l'association du travail de Glennan, Bechtel et Richardson. Premièrement, leurs travaux sont contemporains (1992/1993) alors que ceux de MDC sont ultérieurs (2000). Deuxièmement, la caractérisation sommaire de la notion de mécanisme de Bechtel et Richardson s'accorde avec la définition explicite de Glennan : ils conceptualisent cette notion à l'aide de la théorie des systèmes complexes – ce qui n'est pas le cas de MDC. Troisièmement, la première génération de néo-mécanistes soutient que l'explication mécaniste est globale ou générale (elle s'applique aux sciences biologiques et aux sciences sociales) alors que

tente de défendre le mécanisme contre les objections affirmant qu'être mécaniste est une forme d'anti-humanisme (<http://mechanism.ucsd.edu/~bill/> : Mechanism's Bugaboos: Freewill, Values, and Human Dignity).

MDC ne défendent qu'une portée locale aux explications mécanistes (biologie moléculaires et neurosciences). Conséquemment, notre présentation du travail de ces trois philosophes représente la première théorie néo-mécaniste développée ou exhaustive.

D'ailleurs, nous montrerons au troisième chapitre que la seconde théorie néo-mécaniste, développée par MDC, s'appuie tout en problématisant sur le travail de Bechtel, Richardson et Glennan – mais surtout par rapport à ce dernier. MDC lui objecteront 1) la trop grande généralité de sa définition du concept de mécanisme, 2) l'importance centrale de la notion de loi dans son concept de mécanisme⁹⁸ et 3) l'inapplicabilité de celui-ci en neurosciences.

I. Analyse du concept de mécanisme (Glennan)

La contribution de Glennan au mécanisme s'inscrit dans une perspective philosophique classique et ambitieuse. Son objectif premier est de proposer une théorie de la causalité qui repose sur la notion de mécanisme : « [...] a mechanical theory of causation suggests that two events are causally connected when and only when there is mechanism connecting them. »⁹⁹. Il essaie, d'autre part, de parvenir à concilier le travail de W. Salmon et de P. Kitcher en défendant la thèse que le mécanisme mène à l'unification des patterns explicatifs¹⁰⁰. Enfin, pour Glennan, le mécanisme a une portée si vaste¹⁰¹ qu'il constitue la théorie de l'explication la plus répandue dans l'ensemble des sciences.

En ce qui nous concerne, nous allons nous concentrer sur la tâche principale du travail précurseur de Glennan, c'est-à-dire le remplacement de la métaphore de la

⁹⁸ Rappelons que le mécanisme est, d'abord, une théorie informelle de l'explication qui dénonce une théorie uniquement formaliste de l'explication (chapitre 1), d'où la réticence de MDC à accorder une trop grande importance à la notion de loi.

⁹⁹ Glennan (1996) *Mechanisms and The Nature of Causation*, p. 64.

¹⁰⁰ Glennan (2002) *Rethinking Mechanistic Explanation*, p. S342 et S343.

¹⁰¹ « Definition (M) [Mécanisme] is meant equally to apply to chemical, biological, psychological and other higher-level mechanisms. » in Glennan (1992) *Mechanisms, Models, and Causation*, p. 31.

machine¹⁰² par une caractérisation explicite du concept de mécanisme¹⁰³. Ainsi, le travail de Glennan possède une force normative que n'a pas l'approche descriptiviste de Bechtel et Richardson (1993)¹⁰⁴ : en fixant une définition du concept même de mécanisme, toute entité qui est exclue par la définition ne peut être sujette à une explication mécaniste. D'ailleurs, cet aspect normatif sera revendiqué par Glennan pour motiver l'exclusion de la physique quantique du domaine de ce qui est mécanistiquement explicable.

Son projet s'articule sur trois modes d'analyse. Nous rapporterons dans la première partie ses positions ontologiques sur le concept de mécanisme : un mécanisme est un système complexe dont l'interaction immédiate entre les parties produit un comportement *x*. Dans la deuxième partie sera exposé le pendant épistémologique du concept de mécanisme : la représentation des mécanismes s'effectue par des modèles mécanistes, un modèle étant un exemplaire kuhnien. La troisième partie développe la théorie mécaniste de la causalité mais nous ne l'aborderons pas dans ce mémoire : nous nous contenterons de souligner qu'une relation causale est causale en vertu du fait que les entités qui la compose sont liées de façon mécaniste.

¹⁰² *Ibid.*, p. 4 : « Mechanism involves commitment to the principle that the world is like a machine and that natural phenomena should be explained in something like the way in which behavior of machines is explained. ».

¹⁰³ *Ibid.*, p. 22.

¹⁰⁴ Bechtel et Richardson, en 1993, n'ont pas encore caractérisé explicitement le concept de mécanisme : ils se contentent d'affirmer que le mécanisme est un système complexe – mais nous y reviendrons dans l'introduction de la deuxième partie de ce chapitre. De plus, nous le constaterons, l'objectif de leur travail est d'illustrer que la recherche scientifique, dans les sciences de la vie, suit un certain *pattern*, qui consiste à décomposer les systèmes et à localiser clairement ses composants. Leur mandat est donc de produire une synthèse descriptive de la recherche scientifique. L'unique forme de normativité que l'on peut trouver dans leurs travaux, à ce qu'il nous semble, est qu'ils s'appuient sur des découvertes scientifiques faisant autorité. Ainsi, ils ont produit une synthèse descriptive de certaines grandes découvertes de la recherche scientifique.

1) L'ontologie mécaniste

Glennan conçoit son travail sur le mécanisme comme une entreprise de clarification qui prend la forme d'une définition explicite de ce qu'est un mécanisme. Le mandat d'une telle entreprise est d'abord d'ordre ontologique puisqu'elle vise à décrire une classe d'entité dans le monde¹⁰⁵. Comme nous l'avons déjà souligné, Glennan affirme que la classe des entités devant être reconnue comme des mécanismes est très vaste puisqu'elle couvre tous les domaines des sciences à l'exception de certaines parties de la physique (où les lois sont considérées comme « fondamentales », c'est-à-dire qu'elles ne peuvent pas être expliquées par d'autres lois : la loi de la gravitation universelle et les équations de Maxwell en sont des exemples).

La définition du concept de mécanisme de Glennan est la suivante :

A mechanism for *x* is a complex system whose behavior *x* is the result of the interaction of a number of parts according to laws (either fundamental or mechanically explicable) of immediate interaction.¹⁰⁶

Avant d'aborder la clarification de la définition du concept de mécanisme, trois remarques préliminaires s'imposent. Puisque un même mécanisme peut servir à produire plusieurs fonctions, les mécanismes sont polyfonctionnels. Ce fait rend nécessaire l'identification fonctionnelle d'un mécanisme lors d'une analyse : « Notice that (M) is a definition of a 'mechanism for *x*' rather than a mechanism *simpliciter*. It is essential that a mechanism does some particular sort of thing, and everything about

¹⁰⁵ *Ibid.*, p. 23.

¹⁰⁶ *Ibid.*, p. 24. Les autres définitions sont celles-ci :

« A mechanism underlying a behavior is a complex system which produces that behavior of the interaction of a number of parts according to direct causal laws. », in Glennan (1996), p. 52.

« A mechanism for *x* a behavior is a complex system that produces that behavior by the interaction of a number of parts, where the interactions between parts can be characterized by direct, invariant, change-relating generalizations. », in Glennan (2002), p. S344.

it, even what *it* is, depend upon what it is doing. »¹⁰⁷. La dernière remarque est une conséquence des deux premières : les mécanismes sont toujours les mécanismes d'une fonction mais ils sont polyfonctionnels, c'est donc l'agent qui doit fixer une certaine fonction au mécanisme. Cette imposition d'un cadre fonctionnel au mécanisme place irrémédiablement les explications mécanistes dans une perspective pragmatique¹⁰⁸.

Trois éléments déterminent la compréhension de la définition : « partie », « loi d'action immédiate » et « loi fondamentale ou mécanistiquement explicable »¹⁰⁹. Nous allons expliciter ci-dessous ces notions.

Partie

Définir ce qu'est une partie d'un mécanisme revient à fixer l'extension explicative d'une approche mécaniste de l'explication : en délimitant le domaine des entités acceptables à titre de partie, on limite le champs d'application du mécanisme. Le défi pour Glennan est celui de toute définition extensionnellement importante mais qui se veut signifiante : une grande généralité s'accompagne d'une perte de compréhension. En effet, la définition consiste à limiter un ensemble M d'entités qui possèdent toutes un nombre n de propriétés. Il faut s'attendre à ce que le définiens d'un définiendum très vaste soit moins bien caractérisé (le n est inférieur) qu'un définiendum très restreint : le définiens du définiendum « Aristote » est largement plus explicite que le définiens du définiendum « entité ». En ce qui concerne le problème de définir ce qu'est une partie de mécanisme, le danger réside en ce que le définiendum est vaste (il couvre le terme mécanisme de toutes les sciences à l'exception de certaines parties de la physique) et donc il est difficile d'avoir un nombre de propriétés spécifiques

¹⁰⁷ Glennan (1992) *Mechanisms, Models, and Causation*, p. 24.

¹⁰⁸ Comme nous l'avons vu dans le premier chapitre, le pragmatisme peut être parfaitement compatible avec le réalisme. Ces trois remarques sont partagées par l'ensemble des néo-mécanistes, nous le constaterons pour MDC dans le chapitre 3.

¹⁰⁹ Glennan utilise de façon équivalente les expressions « laws (either fundamental or mechanically explicable) of immediate interaction » et « direct causal laws ».

suffisant – pour ne pas que le terme « mécanisme » soit aussi vide que celui « d'entité » – sans toutefois être trop limitatif (si l'on souhaite pouvoir l'appliquer à une grande échelle).

La stratégie de Glennan est d'opter pour une définition par la négative en imposant deux restrictions : toute entité conforme à ces restrictions peut être déclarée une partie d'un mécanisme. La première est destinée à éviter deux possibilités : « Parts of mechanisms must be *physical objects*. [...] it is meant to prevent arbitrary properties of objects from being parts; and second it is meant to require parts to be more than just artifacts of a particular description of a mechanism. »¹¹⁰. Il s'agit simplement d'exclure les objets abstraits et non physiques (exemple de Glennan : les nombres et les propositions mais, dans une perspective plus historique, on peut penser aux propriétés attribuées par Aristote, et ses émules scholastiques, à la matière par la doctrine des formes et des qualités que le mécanisme a farouchement combattu – et éliminé¹¹¹). Toutefois, la structure interne des parties peut être simple ou complexe, localisée ou distribuée et n'importe quelle entité qui « survient » sur des entités physiques est une entité physique.

La deuxième restriction imposée sur la notion de partie est que celle-ci soit robuste. Dans le cadre de la recherche scientifique, la robustesse est plus que la convergence sensorielle intermodale puisque qu'elle inclut aussi la convergence entre les instruments utilisés : « It is particularly important that objects be robust in the face of experimental manipulation. »¹¹². Les parties sont donc des objets physiques que nous distinguons des artefacts par le critère de la manipulabilité. Les objets authentiques peuvent être utilisés, grâce à leurs propriétés causales, pour opérer des expérimentations. Glennan défend donc la thèse du réalisme expérimental déjà

¹¹⁰ Glennan (1992) *Mechanisms, Models, and Causation*, p. 31.

¹¹¹ Boas (1956) *The Establishment of the Mechanical Philosophy*, p. 415 : « His [Aristote] four elements, earth, air, fire and water, though all composed of the same universal matter and capable of being converted one into the other, were differentiated by the associated "qualities" of cold, dry, hot and moist. When the elements combined to form more complex substances the qualities which were innate in the matter composing the elements also combined, producing the properties of the mixture. ».

¹¹² Glennan (1992) *Mechanisms, Models, and Causation*, p. 33.

présentée par Hacking (1989) : « La preuve « directe » de l'existence des électrons, par exemple, est notre capacité à les manipuler en utilisant certaines propriétés causales bien comprises. »¹¹³.

Imposer à la notion de partie d'être un objet physique robuste a une conséquence importante au sein des domaines admissibles aux explications mécanistes car il existe des cas où la décomposition des parties mécanistes en objet est impossible. De tels cas doivent être exclus de l'ensemble des phénomènes mécanistiquement explicables. Les champs électromagnétiques offrent un bel exemple. Il est possible d'expliquer les propriétés d'un champ par l'interaction de deux autres champs, ce qui semble correspondre à une explication mécaniste. Toutefois, l'analyse subséquente des deux champs ne peut être effectuée : les champs ne sont pas constitués d'objets physiques robustes.

It is crucial to a mechanical account that a system display some behavior which can be explained by reference to underlying properties of its constituents. But there are not, in this case, constituents with underlying properties. As Hertz remarked, "Maxwell's theory is Maxwell's equations".¹¹⁴

Finalement, tout système dont les parties ne sont pas des objets physiques robustes ne peut être un mécanisme et, *ipso facto*, servir d'explication mécaniste à un comportement *x*.

Interaction directe

Les interactions entre les parties sont généralement appréhendées par des lois, c'est-à-dire des généralisations (proposition) qui supportent les contrefactuels. Certaines propositions générales contrefactuelles peuvent être des généralisations accidentelles, ce qui signifie que les parties peuvent être simplement liées entre elles corrélativement plutôt que causalement (l'évènement *C* cause deux événements *A* et

¹¹³ Hacking (1989) *Concevoir et expérimenter*, p. 437.

¹¹⁴ Glennan (1992) *Mechanisms, Models, and Causation*, p. 35

B sans que A et B soient liés causalement : c'est le fameux problème des causes communes). Afin de remédier à ce problème, Glennan soutient que les parties des mécanismes interagissent de façon directe ou immédiate. Considérons l'exemple suivant : trois roues dentées R_1 , R_2 et R_3 .

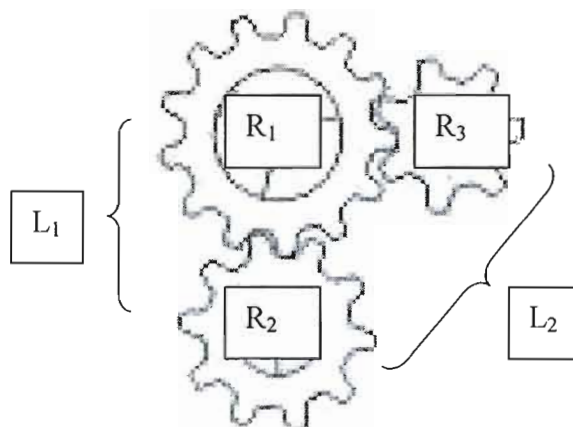


Fig. 5 : l'immédiacité mécanistique

La loi L_1 décrit la rotation de R_1 en fonction de celle de R_2 alors que la loi L_2 décrit la rotation de R_3 en fonction de la rotation de R_2 . L_1 est une loi qui caractérise une interaction immédiate alors que L_2 n'est pas une interaction immédiate puisque la relation entre R_3 et R_2 est médiatisée par R_1 . Dans le cas des roues dentées, la notion de contact direct s'applique mais Glennan opte pour l'expression « Mechanical Immediacy (MI) » pour décrire un contact direct qui n'est pas limité au mécanisme stricte mais qui épouse plutôt le mécanisme méthodologique (qui n'est pas limité a priori à un nombre défini de forces). La notion « d'immédiacité mécanistique » n'est pas équivalente à la notion de contact direct des mécanistes du XVII^{ème} siècle en ce que le contact direct est limité à la seule force mécanique, alors que l'immédiacité mécaniste n'impose aucune limite a priori sur les types de force existants et reconnaît actuellement, en conformité avec le corps scientifique, quatre types de force. Pour rappeler le fait que le néo-mécanisme ne se limite pas aux forces « mécaniques », il serait possible de décrire un mécanisme de trois transformateurs liés entre eux de

façon analogue aux roues dentées et de conclure qu'il existe une immédiacité mécanistique entre eux – excepté pour la relation médiatisée entre R_2 et R_3 par R_1 . Ainsi, pour Glennan, les parties de tout mécanisme sont immédiates mécanistiquement et sont susceptibles d'être caractérisées par des lois d'action immédiate¹¹⁵.

Finalement, les dernières notions à définir sont les lois fondamentales et les lois mécanistiquement explicables. Bien que ces deux types de lois peuvent servir à décrire l'interaction entre des parties d'un mécanisme, il est intéressant d'exposer ce qui les différencie car cette distinction a comme conséquence l'exclusion de la microphysique du champs des phénomènes mécanistiquement explicables – ce qui constitue, comme nous l'avons répété, une marginalisation sans précédent d'une partie importante de la physique de tous les autres champs du savoir scientifique. Les lois mécanistiquement explicables sont celles dont les mécanismes sous-jacents au comportement permettent de rendre compte de ce comportement alors que les lois fondamentales sont uniquement des formalisations descriptives de la régularité du comportement : l'impossibilité de décomposer plus profondément un mécanisme rend impossible l'explication de la régularité aux moyens de l'interaction des parties. C'est le cas pour les équations de Maxwell ou la loi de la gravitation universelle.

2) L'épistémologie mécaniste

L'épistémologie mécaniste peut être scindée en deux types de questions : comment nous représentons-nous les mécanismes de la nature et comment jugeons-nous de la pertinence de ces représentations ? Ces interrogations recouvrent évidemment le problème de la représentation du monde et de la justification de nos représentations. Glennan défend deux thèses qui offrent une réponse à ces deux

¹¹⁵ Il est important de souligner que la notion d'immédiateté mécanistique n'exclut pas a priori que les parties interagissantes soient spatialement distancées. L'argument défensif de Glennan est que l'on ne peut exclure a priori une force qui s'exercerait à distance (Glennan (1992) *Mechanisms, Models, and Causation*, p. 44).

questions. D'abord, nous nous représentons les mécanismes par des modèles mécanistes, ensuite, nous jugeons de leur pertinence en fonction de leur similitude avec la réalité.

Les modèles mécanistes

Les modèles mécanistes sont définis comme suit :

A mechanical model is a description of a mechanism, which comprises (i) a description of the mechanism's behavior, and (ii) a description of the mechanism which accounts for that behavior.¹¹⁶

Les modèles mécanistes sont donc les supports représentatifs de notre connaissance à propos des mécanismes eux-mêmes. La description du comportement du mécanisme peut être nommée description comportementale ou description externe alors que la description du mécanisme responsable du comportement peut être nommée description mécaniste, description structurale ou description interne. La logique de la distinction est simple : il s'agit de caractériser l'explanans, ce que fait le mécanisme, et pour ensuite fournir l'explanandum, c'est-à-dire montrer comment l'explanans est produit.

Les modèles sont des modèles symboliques et non pas physiques, ils ne sont pas des copies concrètes de systèmes matériels (comme une maquette d'avion est un modèle d'un avion réel). Ils ne sont pas non plus des modèles métaphysiques comme a pu l'être le mécanisme cartésien – rappelons que le mandat principal de Glennan est d'explicitier la notion de mécanisme afin qu'il ne soit plus une vague métaphore du fonctionnement de la nature.

La notion de modèle mécaniste s'apparente à la notion d'exemplaire de Kuhn. Un exemplaire kuhnien est un cas concret de résolution de problème qui sert autant de guide de recherche pour la résolution d'autres énigmes du paradigme que d'outil pédagogique (il montre, par analogie, l'application des généralisations symboliques).

¹¹⁶ Glennan (1992) *Mechanisms, Models, and Causation*, p. 49.

En effet, selon lui, une partie de l'apprentissage des sciences est informelle puisqu'il s'appuie sur la capacité de l'agent à reconnaître comment une nouvelle énigme peut être résolue par une méthode de résolution déjà existante : « Les scientifiques résolvent des énigmes en les modelant sur des solutions précédemment trouvées à d'autres énigmes, souvent avec un recours minimum aux généralisations symboliques. »¹¹⁷. Les modèles mécanistes remplissent les mêmes rôles : étant idéalisés, ils peuvent orienter la résolution de nouvelles énigmes en fonction des exemples et méthodes passées. Toutefois, ces extensions de l'application des modèles à d'autres énigmes peuvent considérablement se complexifier en fonction du nouvel objet d'étude et modifier le modèle original. Les modifications apportées sur le modèle original lorsque confronté à de nouvelles énigmes peuvent être d'amplitude variable (de la simple application jusqu'à la réinterprétation éliminativiste¹¹⁸) et connaître a priori l'importance de ces modifications semble une entreprise vaine. Il faut toutefois reconnaître que Glennan ne possède pas les ressources conceptuelles pour rendre compte de l'élaboration effective des modèles mécanistes en sciences. Affirmer que les résolutions de problèmes s'effectuent par « analogie » est assez commode mais quelque peu superficiel : mais ce problème prégnant de l'élaboration des modèles explicatifs sera justement abordé avec les heuristiques de Bechtel et Richardson.

Malgré tout, l'idée maîtresse des modèles mécanistes est identique à celle qui soutient les heuristiques de Bechtel et Richardson : la dynamique de la recherche scientifique est celle d'un système cognitif qui doit solutionner un problème sous de multiples contraintes. Aux limitations intrinsèques du système cognitif, s'ajoute deux contraintes externes. Les artefacts représentationnels (exemplaires, généralisation symbolique, etc.) qui permettent de modéliser de nouvelles énigmes analogiquement

¹¹⁷ Kuhn (1983) *La structure des révolutions scientifiques*, p. 258.

¹¹⁸ Il s'agit d'une des possibilités méthodologiques des heuristiques de Bechtel et Richardson. Lorsqu'il faut décomposer un sous-système qui contrôle une certaine fonction (dans un système) et localiser ses composants, il se peut que le niveau descriptif adopté ne soit pas adéquat : il faut alors reconceptualiser le phénomène à un autre niveau descriptif. Nous reviendrons sur ce cas dans la prochaine section.

sont une ressource limitée. La deuxième contrainte externe est la conformité de nos modèles à ce que croyons être la réalité : ce qui soulève l'épineuse question des critères permettant de déclarer un modèle « similaire à la réalité ».

Théorie de la similarité

Affirmer que nous jugeons de la pertinence du modèle d'un mécanisme en fonction de leur similitude avec la réalité équivaut à dire que l'explication est une simulation fidèle (ou relativement fidèle) de la réalité. Selon Glennan, offrir un modèle mécaniste d'un comportement *x* et affirmer sa similarité avec le système concerné constitue une explication¹¹⁹.

Il est difficile de soutenir une telle théorie de l'explication sans soulever de multiples objections en provenance du débat réalisme/anti-réalisme et, de façon plus compromettante, d'une théorie correspondantiste de la vérité. Pourtant Glennan réclame une totale neutralité de sa théorie envers les enjeux susnommés¹²⁰. Sa stratégie consiste à soutenir que les considérations pragmatiques inhérentes aux explications scientifiques permettront de trancher la question de la similitude sans faire appel à une théorie philosophique de la vérité et à une position dans le débat réalisme/anti-réalisme¹²¹.

L'idée sous-jacente à cette prise de position est que les considérations pragmatiques contraignent suffisamment les agents pour qu'ils s'accordent entre eux sur le degré acceptable de similitude entre un modèle et un mécanisme sans une formalisation théorique générale des rapports représentation-réalité. À la limite, on pourrait argumenter qu'advenant des divergences entre scientifiques concernant le

¹¹⁹ « Haugeland (1978) even suggests that certain entities which philosophers call explanations are what scientists refer to as 'models'. », *ibid.*, p. 58.

¹²⁰ *Ibid.*, p. 51.

¹²¹ Glennan croit que ce débat n'a rien de substantiel : les réquisits à posséder pour toute théorie de la relation représentation-réalité sont la possibilité d'expliquer la faillibilité et la reconnaissance de la finitude de notre accès réalité. D'après Glennan, les versions sophistiquées des deux camps sont en mesure de satisfaire ces réquisits, *ibid.*, p. 9.

degré de similitude entre deux modèles d'un même mécanisme en dépit des contraintes pragmatiques, la communauté scientifique ne s'en porterait pas plus mal : peut-être que la multiplication des expérimentations permettra de trancher a posteriori la question. En effet, il est illusoire de penser qu'il n'existe pas une cohabitation continuelle de points de vue opposés, ou plus simplement divergents, entre les divers acteurs de la communauté scientifique. Il faut considérer que l'indécidabilité entre deux modèles concurrents, quant à savoir lequel est le plus conforme à la réalité, comme un phénomène courant. D'ailleurs, bien souvent, il faut recourir à des analyses comparatives et tenter de combiner les avantages de diverses approches.

Une première contrainte pragmatique à l'évaluation d'un modèle et la spécification du niveau de description à partir duquel le modèle est construit (ce qui n'implique pas que tous les niveaux sont également pertinents pour décrire un phénomène). Par exemple, le mécanisme de la dépolarisation de la membrane peut être décrit dans le contexte plus large du mécanisme de transmission nerveuse ou dans le contexte du mécanisme de la mémoire (conçue comme un renforcement inter-synaptique). Le rôle causal de la dépolarisation de la membrane n'est pas équivalent dans les deux contextes. Pour comprendre la transmission synaptique, le mécanisme de dépolarisation suffit ; pour comprendre le phénomène de la mémorisation, le mécanisme ne suffit plus puisqu'il faut introduire de multiples éléments externes (le système glutamatergique). Ainsi, le niveau descriptif à partir duquel nous analysons un mécanisme est un facteur pragmatique qui peut changer les critères selon lesquels un modèle mécaniste est jugé similaire à la réalité. En bref, puisque les impératifs de rigueur de la recherche scientifique ont pour conséquences le choix d'une perspective particulière sur l'étude d'un mécanisme, évaluer la similarité du modèle par rapport au mécanisme réel doit s'exécuter dans le cadre de cette perspective.

Ensuite, la contextualisation de la description est une contrainte essentielle à l'évaluation des modèles : la contextualisation consiste à considérer les propriétés de système en fonction de l'environnement versus la décontextualisation qui considère plutôt les propriétés du système comme étant intrinsèques – c'est ce qui distingue

l'étude évolutionniste des organismes de l'étude physiologique des organismes. Par exemple, une montre peut être analysée dans le contexte d'un cours d'ingénierie mécanique pour son organisation matérielle mais peu aussi être analysée dans un contexte social interprétatif comme mécanisme signifiant l'heure (contexte non interprétatif vs contexte interprétatif).

Une autre série importante de contraintes pragmatiques comprend un certain nombre de présupposés implicites et des pratiques admises dans le cadre de chaque discipline scientifique. La littérature en philosophie des sciences réfère parfois à cette série comme l'ensemble des clauses *ceteris paribus*. Un exemple courant de ces clauses sont les conditions de dysfonctionnements des mécanismes (*breakdown conditions*) : les spécifications des conditions normales d'utilisation sont si étendues qu'il est impossible d'en faire la recension exhaustive bien que ces spécifications peuvent être invoquées pour expliquer une anomalie. Devenir familier avec cet ensemble tacite de clauses est une partie importante et informelle (c'est la dimension « ésotérique » de la science, selon l'expression de Kuhn) de l'apprentissage scientifique.

Toutefois, la théorie de la similarité de Glennan mériterait de plus amples développements : il semble quelque peu expéditif de rejeter toute approche formelle de la similarité au profit d'une approche exclusivement pragmatiste. D'ailleurs, plusieurs travaux ont développé des notions d'analogie pour les systèmes d'apprentissage artificiels¹²². De plus, plusieurs chercheurs tentant d'appliquer le mécanisme aux sciences sociales, par l'intermédiaire de simulations informatiques de microsociétés et s'intéressent de près, formellement, aux rapports d'analogie entre leurs simulations et les véritables systèmes sociaux (Sawyer, 2003). Pour finir sur le thème de la similarité, nous verrons au troisième chapitre comment Bechtel et Abrahamsen (2005b) soutiennent un argument en faveur du mécanisme basé sur une théorie de l'esprit et de la représentation. Aussi, nous verrons plus précisément

¹²² Hesse (2000) *Models and Analogies*, in Newton-Smith (2000) *A Companion to the Philosophy of Science*, p. 307.

comment MDC ont fait de la contextualisation un des modes explicatifs du néo-mécanisme (les explications contextuelles ou fonctionnelles) où des aspects formels sont intégrés (par des fonctions mathématiques).

En résumé, Glennan vise le remplacement de la métaphore de la machine par une caractérisation explicite du concept de mécanisme sous trois modes d'analyse. Ontologiquement, un mécanisme est un système complexe dont l'interaction immédiate entre les parties produit un comportement x . Épistémologiquement, la représentation des mécanismes s'effectue par des modèles mécanistes, un modèle étant un ensemble de description mécaniste jouant un rôle semblable aux exemplaires kuhniens (instance idéalisée d'une explication type et donc un outil pédagogique). Un modèle explicatif est un modèle approximativement isomorphe au mécanisme dont il est le modèle. La similitude est évaluée dans le cadre de l'activité scientifique en fonction de multiples contraintes théoriques, techniques et pragmatiques : le besoin d'une définition philosophique en terme de critères formels d'acceptabilité (pour que x soit similaire à y) est donc exclu.

II. L'explication mécaniste comme stratégie de recherche scientifique

L'apport de Bechtel et Richardson concernant le mécanisme est une extension des thèses de Simon et Wimsatt que nous avons exposées plus tôt. Un naturalisme pragmatiste : l'objectif théorique recherché est de rationaliser le processus cognitif de la recherche scientifique, de mettre à jour les contraintes¹²³ sur le raisonnement conçu comme un exercice de résolution de problème. Un engagement réaliste : les explications scientifiques doivent se conformer aux caractéristiques ontologiques des systèmes (la nature et le degré d'intégration de leur composant).

¹²³ Les contraintes sont de quatre ordres : psychologique, phénoménologique, physique et opérationnel, Bechtel et Richardson (1993) *Discovering Complexity*, p. 235.

Précisément, Bechtel et Richardson tentent de cerner ce qu'est une explication mécaniste et comment elles sont produites en biologie, biochimie, psychologie et en sciences cognitives. Selon leur définition, un mécanisme est un système complexe¹²⁴ et une explication mécaniste consiste en l'élucidation du comportement d'un système à partir des fonctions exercées par les parties et les interactions entre les parties¹²⁵. L'activité scientifique d'expliquer mécanistiquement consiste principalement en la localisation et la décomposition des systèmes naturels (ou artificiels) : « A major part of developing a mechanistic explanation is simply to determine what the components of a system are and what they do. »¹²⁶. Dès lors, leur projet prendra la forme d'une famille d'heuristiques, de principes guidant la recherche, qui est utilisée par les chercheurs afin de trouver des solutions adéquates dans l'espace des solutions possibles. La pertinence de leur travail, malgré sa singularité dans le paysage philosophique, est importante dans la mesure où il constitue une étude réaliste des modes de production des explications mécanistes par le milieu scientifique de la biologie et de la psychologie¹²⁷. Soulignons cependant que Bechtel et Richardson s'inscrivent dans les traces de Wimsatt (1980, 1996) qui a lui-même proposé de multiples heuristiques ainsi que des méta-heuristiques. Puisque les heuristiques ne couvrent qu'un aspect des explications mécanistes et pour des raisons d'espace, nous ne les aborderons pas.

L'heuristique de la décomposition et de la localisation a été établie à partir d'un ensemble de recherches et de découvertes scientifiques tirées de l'histoire des sciences. Pour cette raison, Bechtel et Richardson affirment que leur guide de recherche s'applique à ces domaines (la pertinence de leur travail doit donc être évaluée dans ces limites).

¹²⁴ *Ibid.*, p. 18.

¹²⁵ *Ibid.*, p. 17.

¹²⁶ *Ibid.*, p. 18.

¹²⁷ « [...] the pursuit of mechanistic explanations of phenomena has remained a staple of both biology and psychology. », Bechtel et Abrahamsen (2005a), p. 2.

Une dernière remarque générale : bien que leur approche théorique est particulièrement descriptive, Bechtel et Richardson croit qu'elle ne constitue pas un rejet de toute contribution normative pour les sciences. Ils soutiennent plutôt une reconstruction de l'entreprise normative dans la droite ligne de Simon : ils tentent une reconstruction réaliste de la rationalité humaine dans le cadre d'une théorie réaliste de la rationalité scientifique¹²⁸ : « [...] it is fruitless to prescribe what we cannot perform. »¹²⁹.

Nous exposerons en détails deux stratégies de décomposition, la méthode analytique (inspirée de Condillac) et la stratégie synthétique (tirée du programme fonctionnaliste de la philosophie de l'esprit ; Cummins (1984), Dennett (1978)) ainsi que deux types de localisation, la localisation simple et complexe. Ensuite, nous montrerons comment l'heuristique doit s'accorder au type de système étudié en vertu de la dépendance du degré de décomposabilité d'un système envers le degré d'interaction des parties de ce système. En d'autres mots, plus les parties d'un système interagissent, moins il est décomposable et plus l'heuristique est complexe.

1) L'explication mécaniste comme heuristique de décomposition et localisation

Puisque les explications mécanistes exigent que le comportement du système étudié soit une production de l'ensemble coordonné des composants, toute étude mécaniste est essentiellement une tentative de décomposition des systèmes et de localisation de ses composants. Montrer comment une telle entreprise est possible est le mandat des heuristiques de recherche de Bechtel et Richardson.

La décomposition s'occupe de déterminer les composants du système et leurs rôles dans la perspective d'une compréhension d'une capacité ou d'un comportement particulier. Deux stratégies générales sont possibles : la méthode analytique (MA) et la stratégie synthétique (SS). La MA est une décomposition matérielle car elle tente

¹²⁸ *Ibid.*, p. 10.

¹²⁹ *Ibid.*, p. 11.

d'isoler physiquement les composants du système (en suivant les délimitations naturelles, par exemple¹³⁰) pour ensuite investiguer le rôle de chacun de ceux-ci. Il s'agit d'une stratégie ascendante (*bottom up*) puisqu'elle élabore la décomposition à partir d'une solide base anato-morphologique. Une autre stratégie ascendante, que nous avons abordé dans la partie précédente, est la simulation. Il s'agit de reproduire artificiellement, habituellement sur une base informatique, les éléments principaux du système réel et tenter de reproduire la dynamique de leurs interactions. À l'aide de ce modèle plus maniable, plusieurs opérations pourront être programmées pour évaluer le rôle des éléments dans la dynamique générale du système. Nous avons déjà donné l'exemple des microsociétés, mais d'importants travaux ont aussi été réalisés dans le domaine de la vie artificielle (Hesse, 2000).

La principale difficulté rencontrée par la MA est que les systèmes complexes les plus étudiés ne révèlent pas aisément leurs propres partitions : pour le cerveau, par exemple, il n'existe que très peu d'évidences concernant d'éventuelles « frontières naturelles » de composants. À ce titre, les expérimentateurs ont du rivaliser d'ingéniosité pour faire parler la nature. Il existe deux méthodes analytiques bien connues : les expériences inhibitrices et les expériences stimulatrices. Les premières consistent à rendre inopérant un composant particulier et de comprendre son rôle à la lumière des irrégularités que l'inhibition provoque. Les deuxièmes tentent plutôt d'altérer le fonctionnement normal d'un composant en le stimulant au-delà des limites habituelles de fonctionnement afin de lier le composant avec les effets provoqués (nous reviendrons plus en détails sur les méthodes analytiques en neurosciences au chapitre 3). Retenons que l'avantage de la MA est qu'elle peut-être poursuivie sans une théorie préalable du système étudié mais qu'elle possède le défaut rédhibitoire d'altérer la dynamique propre des systèmes étudiés si bien qu'il est toujours possible que l'objet d'étude soit un artefact expérimental – un objet d'étude artificiellement provoqué par nos interventions.

¹³⁰ Dans le chapitre 1, nous avons donné l'exemple de la décomposabilité du système digestif : bouche, œsophage, estomac et intestins...

La SS est une décomposition fonctionnelle car elle conjecture sur la façon dont le système doit produire un comportement à partir des opérations effectuées par les composants (leur fonction ou leur rôle¹³¹) et ensuite identifier ces composants¹³². C'est une stratégie descendante (*top down*) puisqu'elle tente d'identifier physiquement les composants à la suite d'un modèle du fonctionnement du système. Cette stratégie est parfois considérée comme étant spéculative. Toutefois, bien que pouvant être spéculative, la SS n'en est pas moins une étape de la recherche scientifique : lorsque peu de données sont disponibles, il est acceptable de proposer des modèles spéculatifs. Il existe toutefois des risques inhérents à ce type d'exercice : la théorie de la nutrition construite par Liebig en est un exemple. Ce dernier a construit un modèle du métabolisme animal à partir de la comparaison entre la composition chimique des aliments absorbés et rejetés ainsi que la croyance selon laquelle le métabolisme consiste seulement en la décomposition des aliments. Le peu de données détenues par Liebig s'accordait avec son modèle mais, comme le montra Bernard avec le cycle du glucose, la dernière supposition de Liebig est fausse ; le métabolisme animal synthétise des composés chimiques. Les modèles sont et doivent être soumis au tribunal de l'expérience.

La MA et la SS sont vues par Bechtel et Richardson comme des stratégies complémentaires de recherche¹³³. Les hypothèses de travail, et donc la SS, sont inévitables car elles procurent un cadre interprétatif aux données expérimentales, la MA est inévitable pour confronter et affiner les modèles à la réalité. Ce pluralisme de méthodes est typique de la biologie et surtout des neurosciences. Bechtel (2002b) soutient d'ailleurs que l'usage combiné de méthodes et des techniques diverses procure de la robustesse à l'objet étudié et évacue la possibilité récurrente que cet objet soit un artefact expérimental.

¹³¹ Fonction et rôle sont synonymes : la fonction de x se comprend dans un cadre analytique du système s (qui contient x) : la fonction de x est alors le rôle que lui assigne ce cadre analytique dans l'explication du comportement global de s . Une telle approche de la fonction est défendue par Cummins (1984) et Craver (2002, 2004).

¹³² Bechtel et Richardson (1993) *Discovering Complexity*, p. 18.

¹³³ *Ibid.*, p. 21.

Bien entendu, les praticiens ne s'entendent pas toujours sur l'importance à accorder à l'une ou l'autre méthode. L'argument principal pour soutenir une prédominance de la MA est que l'expérimentation modifie nos définitions et modèles initiaux : y accorder trop d'importance est donc une perte de temps. L'argument inverse s'apparente à celui que les phénoménologues (Stumpf, Hering, Husserl) défendent à propos de l'étude de l'esprit : sans modèle précis et définition stricte, pas moyen de poursuivre une étude rigoureuse puisque l'imprécision de notre investigation nous rend opaque l'objet même de l'étude.

Comme nous l'avons précisé avec le travail de Simon, la décomposition « [...] allows the subdivision of the explanatory task so that the task becomes manageable and the system intelligible. »¹³⁴. Cependant, nous avons aussi souligné les présupposés importants de la décomposition et aussi de la quasi-décomposabilité :

Decomposition assumes that one activity of the whole system is the product of the set of subordinate functions performed in the system. It assumes that there are but a small number of such functions that together result in the behavior we are studying, and that they are minimally interactive. We start with the assumptions that the interaction can be handled additively or perhaps linearly. Whether these assumptions are realistic or not is an open question; indeed, at the outset we often simply don't know.¹³⁵

Bien évidemment, les systèmes biologiques sont en majorité des systèmes où les parties interagissent fortement ou très fortement et de façon non linéaire (modèles cycliques, comme le cycle de Krebs). Nous montrerons, dans la deuxième partie, comment la décomposition, et donc l'heuristique, doit s'accorder au type de système étudié en vertu de la dépendance du degré de décomposabilité d'un système envers le degré d'interaction des parties de ce système. Plus simplement, le type de décomposition est fonction du degré de décomposabilité – qui dépend lui-même du degré d'interaction entre les parties de ce système. L'histoire des sciences nous

¹³⁴ *Ibid.*, p. 23-24.

¹³⁵ *Idem.*

enseigne la difficulté de la décomposition et de la localisation ; l'étude actuelle du cerveau¹³⁶ est le témoin patent.

Les détracteurs contemporains de la localisation sont plutôt rares et plusieurs ne s'opposent qu'à une forme simpliste de localisation et non quant à son impossibilité. Hardcastle et Stewart, par exemple, dans *Localization in the Brain and Other Illusions*, soutiennent la thèse selon laquelle il est impossible d'associer une quelconque fonction à une certaine aire cérébrale puisque « The brain always struggles of compensate for its losses, and always change is occurring. »¹³⁷. Autrement dit, la plasticité cérébrale, la capacité à accomplir une même fonction par plusieurs voies cérébrales, empêche de soutenir la thèse de la localisation. Toutefois, dans le même article, les auteurs exposent des modèles de compensation, c'est-à-dire des schémas qui illustrent comment le cerveau peut effectuer le transfert d'une fonction d'une aire à une autre. Il est clair que les auteurs refusent une certaine thèse de la localisation – celle qui affirmerait que la localisation est l'assignation rigide d'une fonction x à une aire y – mais, en revanche, ils soutiennent quand même une certaine thèse localisationniste – celle qui affirmerait qu'une fonction est toujours localisée dans une aire, mais que plusieurs autres aires peuvent remplir cette même fonction et, de fait, il arrive qu'une fonction soit générée par une autre ère suite à un accident ou une intervention. Ainsi, certains des opposants ne s'objectent qu'à une forme simpliste de localisation et non à des formes plus complexes de localisation.

La localisation consiste en l'identification des différentes activités proposées par l'entreprise de la décomposition avec le comportement ou la capacité des composants spécifiques¹³⁸. Il est possible de distinguer deux types de localisation : simple et complexe. La localisation simple est une hypothèse qui assigne un comportement ou capacité à un seul composant du système : la localisation simple est

¹³⁶ Bechtel prône un certain optimisme quant à la possibilité de décomposer et de localiser les mécanismes cognitifs sur la base de la réussite relative d'une telle entreprise pour les mécanismes fondamentaux de la mémoire et le système visuel (Bechtel 2001, 2002a).

¹³⁷ Hardcastle et Stewart (manuscript) *Localization in the Brain and Other Illusions*, p. 13-14.

¹³⁸ Bechtel et Richardson (1993) *Discovering Complexity*, p. 24.

donc la recherche *du* locus qui contrôle l'activité étudiée. Par exemple, s'il on recherche le locus du contrôle génétique dans la cellule, l'unique processus de la décomposition matérielle suffira : le noyau. Ce genre de localisation se retrouve principalement dans les premières ébauches de modélisation scientifique ou lorsque les systèmes étudiés sont relativement simples. L'exemple paradigmatique de la localisation simple en neurosciences est le programme phrénologique de Gall¹³⁹. Ce dernier soutenait la thèse de l'organologie : l'esprit est composé de facultés discrètes localisées (des « organes ») de façon discrète. En bref, l'esprit est un ensemble agrégé d'organes indépendants. La localisation était rendue possible grâce à l'hypothèse craniologique : le crâne est malléable et prend la forme du cerveau, celui-ci est donc révélateur de l'importance relative de chaque « organe » cérébral.

Naturellement, toute localisation simple est simpliste : elle n'en constitue pas moins une étape inaugurale dans l'étude d'un système et peut apporter de précieuses indications heuristiques pour la recherche ultérieure. Gall a tout de même été l'un des premiers pourfendeurs influant de la théorie cartésienne de l'esprit immatériel et a aussi été une source d'inspiration pour les travaux plus sérieux de localisation (complexe) d'un Broca ou d'un Wernicke.

La leçon générale tirée de l'étude de plusieurs tentatives de localisation simple est que celles-ci échouent la plupart du temps, ce qui nous mène à une double possibilité : abandonner le programme mécaniste ou tenter une localisation complexe. La première solution a été celle des vitalistes tels Flourens et Bichat. Bien que de multiples objections au travail des phrénologistes soient légitimes, il n'est pas concevable pour autant d'embrasser une solution aussi radicale que le vitalisme. La solution prise par la communauté scientifique, et qui s'est révélée fructueuse, est d'engager un effort de reconceptualisation du phénomène étudié vers une complexification des modèles¹⁴⁰. Cette complexification s'effectue par une

¹³⁹ Bechtel et Richardson (1993) nous offrent aussi des analyses de localisations simples pour les phénomènes de l'oxydation et de la respiration, p. 74-88.

¹⁴⁰ *Ibid.*, p. 124.

prédominance de la décomposition fonctionnelle¹⁴¹ et l'élaboration d'une localisation complexe.

Une localisation complexe est une « localization multiply constrained; that is it proposes a set of components that contribute differentially to system function, and it incorporates independent constraints on allowable mechanisms from lower levels. »¹⁴². Deux types de contraintes sont possibles pour localiser une fonction ou une sous-fonction : les contraintes descendantes (*top down*) et les contraintes ascendantes (*bottom up*). Pour le domaine des neurosciences, les contraintes descendantes proviennent de la psychologie (parfois de la philosophie) et les contraintes ascendantes proviennent de la neurobiologie. À l'instar de la décomposition, où il est souhaitable de combiner les méthodes MA et SS, il est souhaitable, et même nécessaire, d'après Bechtel et Richardson, de combiner les deux formes de contraintes pour la localisation¹⁴³.

Toutefois, indépendamment de la nature de la décomposition et de la localisation, le réalisme dont sont partisans Bechtel et Richardson implique que toute fonction doit être matérialisée d'une quelconque façon. De multiples outils sont disponibles pour mettre au jour l'intégration matérielle des fonctions. D'ailleurs, cette entreprise est devenue un domaine spécifique d'importance croissante à l'intérieur des neurosciences : la neuroimagerie. Le mandat des « *brain imagers* » est d'identifier les zones cérébrales actives pour une tâche cognitive donnée, il est donc permis d'affirmer que la localisation est devenue une spécialisation à part entière¹⁴⁴.

¹⁴¹ « Frequently a program of research begins with direct localization, which then develops into a more complex localization in which functional decomposition of tasks assumes a more central role. », Bechtel et Richardson (1993) *Discovering Complexity*, p. 125.

¹⁴² *Idem*.

¹⁴³ Cette approche intégrative des disciplines est acceptée chez tous les néo-mécanistes, preuve qu'il existe des réductionnismes nuancés contrairement à ce que Schouten (2005) semble entendre en présentant le réductionnisme par la lorgnette du *New Wave Reductionism* de Bickle (2003). La position de Schouten a pour conséquence de polariser le débat sur le réductionnisme à un débat disciplinaire entre psychologie et neurosciences où la psychologie n'a qu'un rôle secondaire. C'est donc un des mérites du néo-mécanisme que de déplacer le débat du réductionnisme sur le terrain d'une théorie de l'explication (et non d'un stérile débat disciplinaire).

¹⁴⁴ Pour un compte-rendu sur les techniques de neuroimagerie, voir Bechtel (2005).

Pour donner une idée de l'acuité de ces techniques, certains chercheurs (Beauregard 2001, 2003) tentent de localiser les mécanismes cérébraux responsables du contrôle volontaire des émotions. Cette spécialisation de la localisation tient sans doute à un avantage inestimable de certaines techniques d'imagerie : elles sont non invasives – ce qui permet d'étudier un même sujet à différents intervalles temporels sans altérer le fonctionnement interne de l'organisme (les scanners mesurent le *Blood Oxygen Level Dependence*, BOLD, c'est-à-dire qu'ils mesurent l'activité cellulaire par le niveau d'oxygénation du sang).

En résumé, toute étude mécaniste est essentiellement une tentative de décomposition des systèmes et de localisation de ses composants. La décomposition s'occupe de déterminer les composants du système et leurs rôles dans la perspective d'une compréhension d'une capacité ou d'un comportement particulier et la localisation consiste en l'identification matérielle des différentes activités proposées par l'entreprise de décomposition avec le comportement ou la capacité des composants spécifiques. Deux stratégies peuvent être suivies pour décomposer un système : la décomposabilité matérielle (par la MA) et fonctionnelle (par la SS). La localisation s'opère en fonction du type de décomposition : la DA mène généralement à une localisation simple, la DF à une localisation complexe.

2) Heuristique, décomposabilité et système complexe

Nous avons constaté, dans le premier chapitre, qu'une intuition fondamentale des néo-mécanistes est que le néo-mécanisme en tant que théorie de l'explication et méthode d'investigation scientifique s'appuie sur une théorie ontologique des systèmes complexes et hiérarchiques. Précisément, l'ontologie de Bechtel et Richardson consistera en une taxonomie des systèmes où les différents types de systèmes seront couplés avec une méthode de décomposition-location qui leurs sont adaptés. Le principe de classification de la taxonomie est le degré d'interaction entre les composants d'un système : il y aura donc un continuum entre les systèmes n'ayant

aucune interaction interne (agrégat) et les systèmes ayant une forte interaction interne (système intégré). Évidemment, comme nous l'avons déjà spécifié, la vaste majorité des systèmes naturels possèdent des composants fortement intégrés : le cas des agrégats n'est présenté que pour montrer l'importance de l'organisation interne (l'agencement spatio-temporel et la nature des composants) des mécanismes dans la compréhension de la capacité ou fonction étudiée. Comme les chercheurs ne connaissent pas a priori la complexité organisationnelle d'un système, ils procèdent habituellement par une supposition de la plus grande simplicité possible et procèdent ensuite par complexification.

Naturellement, hormis la contrainte ontologique, d'autres contraintes interviennent dans le choix de la bonne stratégie de recherche, comme les contraintes liées à la rationalité limitée des agents cognitifs (capacités de mémorisation et de computation, tradition de recherche mais aussi les contraintes sociales, comme les comités d'éthiques – qui imposent de fortes limitations sur la construction des modèles animaux, etc.).

Le projet de Bechtel et Richardson s'inscrit donc en droite ligne avec celui de Wimsatt, créer une taxonomie des types d'organisation des systèmes, mais y apporte une contribution importante : la possibilité de construire une description générale de l'évolution des recherches en biologie. Cette description comprend un ensemble de points auxquels les chercheurs se butent et auxquels il n'existe que peu de solutions. D'abord, les scientifiques procèdent d'abord par l'isolation du système étudié pour ensuite proposer une décomposition matérielle ou fonctionnelle du système. Lorsque cette étape est complétée, il s'agit de localiser physiquement, de façon simple ou complexe, le siège du locus de contrôle de la fonction étudiée. La troisième étape consiste donc en la décomposition et la localisation du système d'une fonction spécifique.

Brièvement donc, le choix de l'heuristique est fonction de l'organisation des systèmes. En créant une taxonomie des systèmes, il sera possible d'associer une heuristique spécifique pour un type d'organisation spécifique. Trois types de système

agrégatif, composite et intégré sont donc, respectivement, décomposable, quasi-décomposable et minimalement décomposable. L'agrégat est parfaitement décomposable car le comportement du système est constant indépendamment de 1) la substitution des parties, 2) un changement du nombre de partie, 3) de la réagrégation des parties et 4) d'une interaction minimale entre les parties¹⁴⁵. C'est par le quatrième critère que les systèmes composites se distinguent le plus de l'agrégat : dans un système de type composé, l'organisation des parties est importante pour la fonction mais reste secondaire alors que dans un système intégré, l'organisation des parties est primordiale pour comprendre les capacités du tout¹⁴⁶.

Des exemples typiques de cas extrêmes comme l'agrégat et le système intégré sont la pile de livre¹⁴⁷ et les réseaux connexionnistes (Bechtel). Le premier exemple montre qu'en regard de certaines propriétés (le poids d'une pile de livre), l'organisation spatiale des systèmes agrégatifs (une pile de livres) n'est d'aucune importance : le poids des livres ne dépend pas de l'ordre dans lequel ils sont empilés. À l'inverse, le deuxième exemple montre que l'interaction entre les composants (les neurones artificiels) sont si complexes que l'on renonce à expliquer ce type de système de façon mécaniste en privilégiant une approche purement fonctionnaliste – où la relation entre l'input et l'output dudit système est uniquement appréhendée de façon algorithmique.

Un autre exemple des problèmes liés à la grande complexité des interactions dans un système intégré est le débat concernant la possibilité de comprendre mécanistiquement les fonctions cognitives supérieures. Un des premiers opposants au localisationnisme du XIX^{ème}, Flourens, niait la possibilité même de décomposer le cortex en composants en affirmant que le cortex fonctionne comme un tout indivisible. La position de Flourens a été soutenue jusqu'en 1960, à la mort de Karl Lashley, sous le couvert de l'équipotentialité, ou *mass action*, que l'on peut définir

¹⁴⁵ Ces critères sont ceux de Wimsatt (1986).

¹⁴⁶ Bechtel et Richardson (1993) *Discovering Complexity*, p. 26.

¹⁴⁷ L'exemple est de Pierre Poirier.

comme la thèse affirmant que les capacités cognitives supérieures ne peuvent être divisées en unités fonctionnelles car elles forment une unité fonctionnelle¹⁴⁸.

En résumé, à l'opposé de l'agrégat dont les interactions entre les composants sont inexistantes (ou négligeables), les systèmes fortement intégrés possèdent des interactions complexes (réseaux connexionnistes). L'organisation du système est donc le facteur essentiel pour déterminer le degré de décomposabilité d'un système : plus l'organisation du système rentre en ligne de compte moins le système est décomposable. Ainsi, l'agrégat sera décomposable, les systèmes composés seront quasi-décomposables et les systèmes intégrés seront minimalement décomposables. Selon Finger (1994), il existe une tendance historique en neurosciences qui indique clairement une adoption de plus en plus généralisée du mécanisme et donc de la quasi-décomposabilité. Dès lors, en ce qui concerne l'étude des fonctions cognitives, il faut s'attendre à être confronté à des systèmes composites quasi-décomposables ou, au pire, minimalement décomposables.

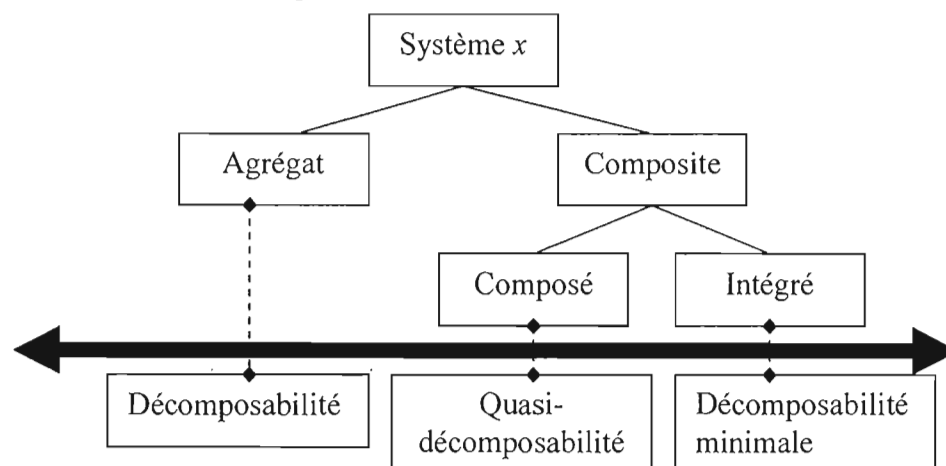


Fig. 6 : taxonomie des systèmes et décomposabilité

Montrons à présent comment Bechtel et Richardson lient cette taxonomie avec le cheminement de la recherche. Les auteurs ont remarqué une certaine régularité

¹⁴⁸ Finger (1994) *Origins of Neuroscience*, p. 35-36 et p. 59-61.

dans le développement de la recherche expérimentale en biologie et psychologie pour les explications mécanistes. Cette régularité s'exprime sous la forme d'une série de points stratégiques où un nombre limité de choix sont possibles – naturellement, ces nœuds névralgiques sont liés à la découverte de la nature du système étudié, donc de la taxonomie exposée ci-dessus.

Le graphique ci-dessous rassemble les multiples étapes qui composent l'heuristique et illustre les « trajectoires possibles » de la décomposition et de la localisation dans le processus de la découverte scientifique. En aucun cas cette heuristique est une garantie de la découverte scientifique ; la thèse de Bechtel et Richardson soutient seulement que de nombreuses découvertes importantes en biochimie, neurosciences, génétique et psychologie cognitive ont suivi de façon variable mais consistante avec cet arbre des possibles.

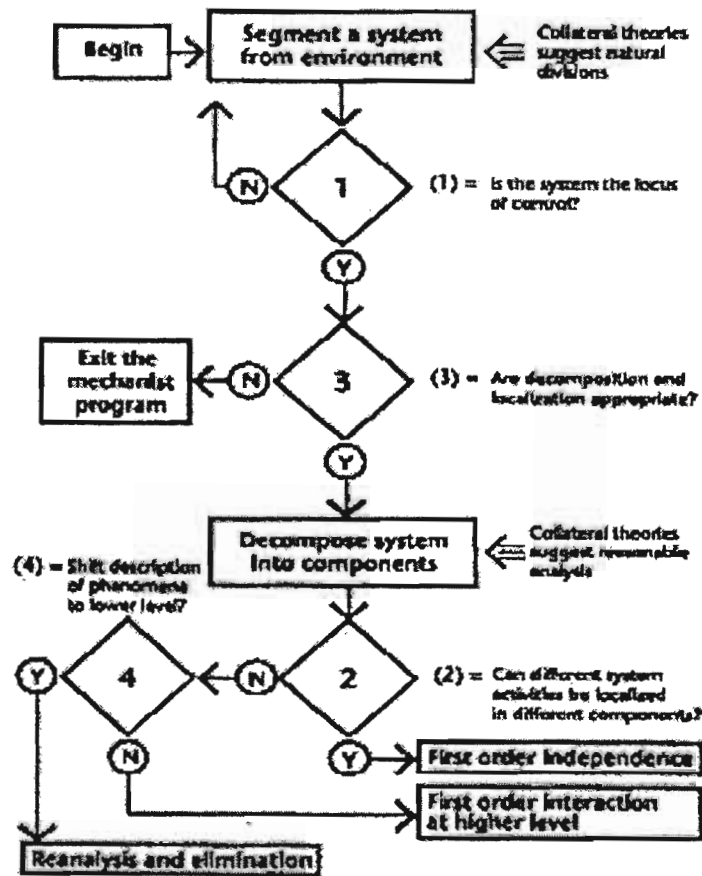


Fig. 7 : heuristique de la décomposition et de la localisation¹⁴⁹

La première étape de la recherche est d'isoler le système de son environnement et d'identifier la fonction à expliquer. Des théories corollaires sont généralement utilisées pour proposer un locus de contrôle de la fonction et les méthodes analytiques sont généralement employées pour corroborer qu'un système physique particulier produit la fonction. Advenant une identification correcte du locus de contrôle, la recherche peut s'affairer à la deuxième étape, la décomposition et la localisation des composants du système responsable du contrôle de la fonction étudiée. Comme nous l'avons déjà souligné, la décomposition et la localisation se modèlent sous la pression de contraintes (conceptuelles et empiriques).

¹⁴⁹ *Ibid.*, p. 195.

La troisième étape consiste donc en la décomposition et la localisation du système en vue de comprendre une fonction spécifique. Cette étape est cruciale : si le système n'est pas décomposable, quasi-décomposable ou minimalement décomposable, l'approche mécaniste est impraticable, ce qui signifie tout simplement qu'il faille abandonner le programme mécaniste. Selon les auteurs, un exemple de « faillite relative » d'analyse mécaniste est le fonctionnement des systèmes connexionnistes (les réseaux de neurones). Comme nous l'avons brièvement remarqué, le programme mécaniste ne peut s'appliquer dans la mesure où il n'est pas possible d'isoler des composants qui sont responsables de la production de sous-fonctions qui, ensemble, produisent la fonction du système. Toutefois, à la défense du mécanisme, il faut remarquer que cette incapacité est une faillite « relative » du programme mécaniste dans la mesure où nous savons que le comportement individuel de chaque neurone artificiel est mécaniste et que le comportement d'un réseau est nécessairement la résultante de l'organisation de l'ensemble des neurones artificielles. Le cas des réseaux connexionnistes est donc un exemple de limite opératoire et non conceptuelle.

Quand il s'agit de comprendre une fonction spécifique au sein d'un système, trois possibilités doivent être envisagées selon que le degré de décomposabilité du système étudié 1) une « indépendance de premier ordre » (si le système est décomposable), 2) une « interaction de premier ordre » (si le système est quasi-décomposable) ou 3) une réinterprétation éliminativiste (si le système est minimalement décomposable).

L'indépendance de premier ordre désigne simplement une localisation simple : « the primary dimensions involved in system behavior can be explained in terms of the behavior of a single component, in light of a functional analysis of the system. »¹⁵⁰. Si cette première étape est un succès, c'est-à-dire qu'il s'avère que le système possède effectivement un locus de contrôle, la recherche pourra s'orienter «

¹⁵⁰ *Ibid.*, p. 147.

verticalement » vers une étude du composant à un niveau inférieur où il faudra à nouveau identifier les sous-composants et effectuer une analyse fonctionnelle du composant. La recherche pourrait aussi s'orienter horizontalement vers l'étude des interactions du locus de contrôle de la fonction avec d'autres composants.

La deuxième possibilité, l'interaction de premier ordre, est que la décomposition mène à une localisation complexe, ce qui signifie que la fonction requiert le concert de plusieurs composants et que le système est seulement minimalement décomposable (c'est un système intégré) : l'organisation et l'interaction entre les composants jouera un rôle important dans le fonctionnement du système et, très souvent, la linéarité est compromise¹⁵¹.

La troisième et dernière possibilité de la décomposition du système, la réinterprétation éliminativiste, est une reconceptualisation du phénomène étudié. L'exemple de Bechtel et Richardson vient de l'histoire de la génétique. La génétique classique (Mendel) prenait pour acquis que les traits phénotypiques étaient sous le contrôle des gènes, unités physiquement discrètes et fonctionnellement autonomes¹⁵². Quelques découvertes comme l'effet de position (T.H. Morgan) et les travaux de Beadle et Tatum ont définitivement ruiné l'adéquation simpliste un trait phénotypique-un gène pour un modèle triadique intégrant le niveau biochimique entre le niveau phénotypique et génétique. Les gènes ne produisent pas des traits phénotypiques mais des protéines qui, elles, sont les composants des traits phénotypiques. Ainsi, au cours du vingtième siècle, la génétique s'est trouvée profondément modifiée : « A characterization of the phenomena in terms of observable traits was replaced by one couched in terms of biochemical products. »¹⁵³. Ce cas est un exemple de reconstitution de la géographie de la connaissance suite à une complexification progressive des interactions entre les acteurs étudiés ; en tant

¹⁵¹ Les auteurs nous offrent le cas de l'étude biochimique de la fermentation cellulaire où Embden et Meyerhof ont contribué à l'abandon d'un modèle linéaire et d'une tentative de localisation simple (Buchner) pour un modèle cyclique ponctué de nombreuses étapes et rétroactions (localisation complexe).

¹⁵² Bechtel et Richardson (1993) *Discovering Complexity*, p. 173.

¹⁵³ *Ibid.*, p. 194.

qu'agent cognitif limité, nous procédons généralement par complexification de nos modèles, complexification qui multiplie les contraintes à imposer – ce qui aboutit éventuellement, après reconceptualisation, à une décomposition et une localisation complexe acceptable des phénomènes étudiés.

Deux remarques nous semblent importantes à propos de cette heuristique. D'abord, elle procède de la supposition la plus simple à propos du système étudié pour progresser graduellement, en cas d'échec, vers des suppositions plus complexes. Dans de nombreux cas tirés de l'histoire des sciences¹⁵⁴, il est pris pour acquis, dans les premiers moments de la recherche, que le système étudié est décomposable ou quasi-décomposable, qu'il est appréhendable de façon linéaire et que la localisation est simple. Généralement, ces présupposés s'avèrent trop simples pour saisir le fonctionnement des entités biologiques et, le plus souvent, les chercheurs doivent multiplier drastiquement les contraintes sur leurs modèles en intégrant des données expérimentales (*bottom up*) et des cadres interprétatifs (*top-down*).

Ensuite, outre cette complexification progressive de l'heuristique de la décomposition et de la localisation, Bechtel et Richardson indiquent certaines limites de l'explication causale mécaniste. Rapidement, ces limites sont de quatre ordres : psychologique, phénoménologique, opérationnelle et physique.

La première limite est une imposition de notre cognition sur le monde externe. À l'instar de Simon, Bechtel et Richardson soutiennent que la cognition n'est possible que lorsqu'elle porte sur des systèmes de type arborescent, ainsi nous prenons pour acquis que les systèmes que nous étudions sont quasi-décomposables. Or le risque qu'ils ne le soient pas est toujours encouru. Toutefois, il pourrait être objecté à Bechtel et Richardson que si l'arborescence est une condition de possibilité à la cognition humaine (c'est l'argument de Simon, voir p. 49), alors tous les systèmes dont nous avons connaissance doivent être arborescents : ainsi, nous ne devrions pas nous inquiéter de la possibilité que nous soyons confrontés à des

¹⁵⁴ Nous avons pu constater que c'est le cas de la neurologie, de la génétique et de certaines recherches pionnières en biochimie.

systèmes non décomposables. Il ne s'agit pas de nier l'existence de tels systèmes, seulement, si Simon a raison – nous ne pouvons appréhender que les systèmes décomposables ou quasi-décomposables – il nous est impossible d'avoir un accès épistémique à ces systèmes. Autrement dit, si la thèse de Simon est vraie, l'ensemble des systèmes non décomposables est simplement inanalysable. Pour finir, cette catégorie de limites comprend aussi les idiosyncrasies culturelles : nous sommes familiers avec l'exemple du mécanisme cartésien qui s'est avéré limité par les sciences de son époque parce qu'elles envisageaient moins de types de force qu'elles ne le font actuellement.

Les contraintes phénoménologiques sont celles imposées par les phénomènes. Un phénomène est habituellement une régularité du comportement d'un système¹⁵⁵. Il est source de contraintes par l'intermédiaire des données que l'on prélève – ce qui entraîne, *ipso facto*, une certaine place à l'interprétation. Nous avons examiné la possibilité que l'heuristique débouche sur une reconceptualisation du phénomène (transformation de la génétique classique en génétique biochimique). Cette contrainte est un corollaire du réalisme de Bechtel et Richardson qui refuse une approche constructiviste de l'objet d'étude (imputée à Hacking : « experiment is the creation of phenomena » 1983, p. 221). La troisième source de contrainte concerne les limites introduites par les instruments et les techniques expérimentales. Fortement occulté en philosophie, le fait que les sciences soient en grande partie un art pratique nous rappelle que la compréhension mécaniste des phénomènes se bute à des problèmes purement pratiques qui dépendent étroitement de l'expérience du praticien – et, parfois, du simple hasard. Pour les systèmes vivants, l'usage des instruments et techniques est prompt à introduire des artefacts expérimentaux, c'est-à-dire la création des phénomènes anormaux causés par l'usage de l'expérimentation. Pour réduire l'introduction d'artefacts, la stratégie répandue est le recoupement de

¹⁵⁵ Bechtel et Richardson (1993) *Discovering Complexity*, p. 238.

techniques : si plusieurs technologies différentes aboutissent à des données similaires, les risques d'introduction d'artefacts sont faibles (Bechtel 2002a).

La dernière contrainte sur les explications causales mécanistes est l'attention accordée à la réalisation physique des fonctions étudiées. En effet, puisque l'explication mécaniste requiert que l'on décompose l'objet d'étude, ce qui signifie que l'on tente de le comprendre à un niveau inférieur, il faut s'attendre à une orientation réductionniste de la recherche. Toute explication est construite à l'aide des composants d'un système, ces composants eux-mêmes compris à l'aide des sous-composants, etc. Le problème est que les informations aux niveaux inférieurs ne sont pas toujours disponibles. Le pouvoir explicatif du mécanisme est donc corrélé avec la masse d'information disponible qui permet de creuser plus profondément dans l'intimité des rouages de la nature.

En conclusion, Bechtel et Richardson ont tenté de caractériser les principales avenues qui se présentent lorsque une explication causale-mécaniste des phénomènes est entreprise. Les scientifiques procèdent d'abord par l'isolation du système étudié pour ensuite proposer une décomposition matérielle ou fonctionnelle du système. Lorsque cette étape est complétée, il s'agit de localiser physiquement, de façon simple ou complexe, le siège du locus de contrôle de la fonction étudié. Nous avons expliqué comment les sciences semblent procéder par complexification en intégrant toujours plus de contraintes (psychologiques, phénoménologiques, opérationnelles et physiques). L'heuristique établie n'a pas valeur de norme et a été construite, nous le rappelons, sur la base d'un échantillon diversifié mais limité de l'histoire des sciences. La leçon principale qu'il faut tirer du travail de Bechtel et Richardson se trouve nuancée : « If we are right, the problem is one of reaching a solution that simultaneously satisfies a complex array of changing constraints. »¹⁵⁶.

¹⁵⁶ Bechtel et Richardson (1993) *Discovering Complexity*, p. 244.

Après avoir dressé, dans le premier chapitre, les thèses sous-jacentes au renouveau de la philosophie mécaniste en le replaçant dans le contexte d'opposition à l'empirisme logique, nous sommes en mesure de comprendre pourquoi le néo-mécanisme se développe selon deux grands axes. Ces axes d'analyse sont largement tributaires du réalisme et du naturalisme pragmatiste des précurseurs du néo-mécanisme en ce qu'ils soutiennent la nécessité de conformer les modèles mécanistes aux mécanismes eux-mêmes (la théorie de la similarité de Glennan est un exemple manifeste) et une approche cognitiviste de la recherche scientifique dans le cadre de la philosophie des sciences. Le premier et principal axe de développement est le besoin d'une définition explicite du concept de mécanisme, trop souvent appréhendé de manière intuitive sous la forme de la métaphore de la machine¹⁵⁷. Le deuxième est incarné par Wimsatt, Bechtel et Richardson et consiste en la construction d'heuristiques de recherche dans le but de caractériser les principales avenues qui se présentent lorsque une explication causale mécaniste des phénomènes est entreprise.

La perspective de Glennan est résolument philosophique en ce qu'elle clarifie explicitement le concept même de mécanisme au niveau ontologique et épistémique : les mécanismes sont des systèmes complexes dont les parties interagissent de façon directe et produisent un comportement *x*. Épistémologiquement, la représentation des mécanismes s'effectue par des modèles mécanistes, un modèle étant un ensemble de description mécaniste jouant un rôle semblable aux exemplaires kuhniens (instance idéalisée d'une explication type et un donc un outil pédagogique). Un modèle explicatif est un modèle approximativement isomorphe au mécanisme dont il est le modèle.

Bechtel et Richardson, quant à eux, nous permettent de comprendre comment sont produites les explications mécanismes en précisant les stratégies de recherche scientifique de la décomposition et de la localisation des systèmes complexes (afin de

¹⁵⁷ Cet axe est déclaré « principal » puisqu'il est la source de quelques frictions entre les différents acteurs néo-mécanistes : c'est une question encore ouverte que de savoir si les différentes définitions du mécanisme présentées par Glennan et MDC sont compatibles.

rendre manifeste la façon dont l'organisation interne des composants produit certaines fonctions). Nous avons montré comment Bechtel et Richardson envisagent l'orientation de la recherche en fonction de la nature des systèmes étudiés : les scientifiques résolvent le fonctionnement des mécanismes en postulant, en premier lieu, que c'est un système décomposable dont le locus de contrôle est discret pour ensuite, en cas d'échec, complexifier les suppositions sur le système jusqu'à ce qu'ils parviennent à une décomposition et une localisation qui permettent l'explication des fonctions étudiées. Ainsi, Bechtel et Richardson s'inscrivent dans la lignée intellectuelle de Wimsatt et Simon en ce qu'ils conçoivent la recherche scientifique comme une activité cognitive de résolution de problèmes et donc une gestion de contraintes multiples dans un espace de solutions possibles.

Nous concluons ce chapitre en soulevant le fait que les travaux respectifs de Glennan, Bechtel et Richardson sont compatibles en ce qu'ils développent chacun une dimension différente du mécanisme tout en partageant une notion systémique du mécanisme. Bechtel et Richardson explorent l'aspect stratégique de la recherche scientifique des explications mécanistes et répondent à la première problématique de notre mémoire – celle de la découverte effective des mécanismes – en montrant comment les mécanismes sont découverts en procédant, par une série d'étapes successives, à la décomposition du système étudié et à la localisation des ses composants.

Glennan, quant à lui, met l'emphasis sur l'aspect ontologique et épistémique de la notion de mécanisme ce qui nous a permis d'éclairer notre deuxième problématique – la nécessité de caractériser de façon claire et explicite la notion même de mécanisme. Nous verrons que plusieurs notions incluses dans cette première notion de mécanisme poseront quelques problèmes pour la seconde génération de néo-mécanistes (MDC). Aussi, plusieurs problèmes ontologiques se poseront quant à la question de savoir quelles entités sont acceptables en tant que composants d'un mécanisme.

Chapitre 3

Seconde théorie exhaustive du néo-mécanisme (Machamer/Darden/Craver)

MDC développent une caractérisation de ce qu'est un mécanisme qui est adéquatement descriptive à l'usage qu'en font les scientifiques en biologie moléculaire et en neurosciences. La pertinence de l'analyse de MDC est donc double : nous renseigner de façon réaliste sur une notion clé de ces sciences¹⁵⁸ et combler un déficit en philosophie des sciences (ils croient que la notion de mécanisme est insuffisamment analysée, tout comme c.f. Schaffner 1993, p. 287).

Toutefois, la thèse de MDC n'affirme pas que le mécanisme est la seule explication possible, ni que la découverte de mécanismes est toujours l'objectif

¹⁵⁸ La thèse selon laquelle les mécanismes sont une notion importante pour comprendre la neurobiologie et la biologie moléculaire est soutenue par Bechtel et Richardson (1993), Brandon (1985), Kauffman (1971), Wimsatt (1972), Burian (1996), Crick (1988) mais aussi Glennan (1992, 1996, 2002), Thagard (1998-2003), Woodward (2002), Bogen (2004) et, récemment, par Schaffner (2004). Outre cet appel à l'autorité, tout un chacun peut se livrer à l'exercice simple, et concluant, de lire les introductions de manuels, ou mieux, les articles spécialisés, pour s'apercevoir que les notions importantes pour les explications en neurobiologie et en biologie moléculaire sont celles de fonction et de mécanisme.

premier des recherches. D'ailleurs, Craver a fait un travail considérable pour rendre compte de l'analyse fonctionnelle dans les sciences biologiques et il montre, dans *Functions and Mechanisms in Contemporary Neuroscience* (2004), qu'une explication neuroscientifique complète engage fonctions et mécanismes.

Tout comme pour Glennan, nous préciserons deux axes qui fixent les questionnements philosophiques entourant le mécanisme : 1) l'analyse ontologique du mécanisme (continuité productive, entité, activité et contrainte spatio-temporelle) et 2) les caractéristiques épistémiques du mécanisme (schéma, sketch, niveau et hiérarchie mécaniste, les principes expérimentaux et la théorie causale-mécaniste de l'explication).

Ce dernier chapitre mettra en lumière une deuxième conceptualisation non-systémique du concept de mécanisme et ouvrira une série de discussion au sujet de problèmes propres au champ du néo-mécanisme. Nous verrons comment les lois d'interaction directe de Glennan ont été, maladroitement selon nous, critiquées par MDC : nous monterons qu'ils défendent, sous une autre terminologie, une idée similaire. Ensuite, nous aborderons une autre opposition entre les deux générations de néo-mécanistes concernant l'ontologie : les interrelations sont-elles réductibles à des dispositions d'objets ou sont-elles des entités ontologiques distinctes des objets ?

Finalement, nous nous pencherons plus précisément sur la pratique effective des scientifiques (en biologie moléculaire, neurobiologie et neurosciences) en explicitant en quoi consiste une approche informelle des sciences. Pour ce faire, nous verrons de quelle manière l'étude des propriétés des entités et des activités (les contraintes spatiotemporelles) sont centrales à la construction de nos représentations des mécanismes (par les schémas et les sketches mécanistes).

I. Analyse ontologique du mécanisme

La caractérisation générale du mécanisme de MDC est la suivante :

Mechanisms are entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions.¹⁵⁹

Comme nous le verrons dans l'approche épistémique, les mécanismes sont très souvent représentés par des diagrammes. Celui ci-dessous illustre le mécanisme de la transmission chimique interneuronale. Il nous servira de référence lors de notre caractérisation ontologique du mécanisme.

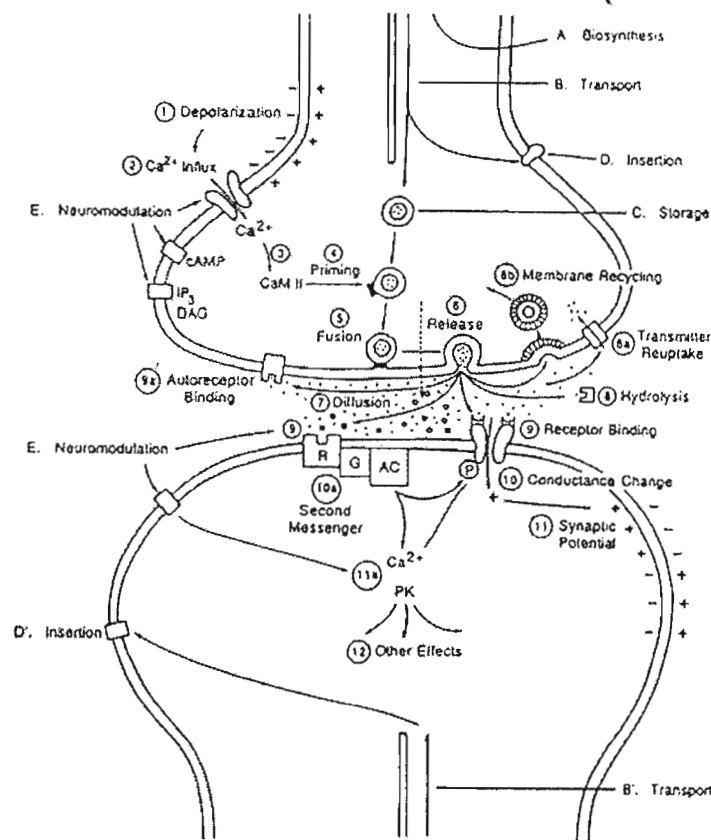


Fig. 4.8 A summary of some of the main biochemical mechanisms that have been identified at chemical synapses. A–E. Long-term steps in synthesis, transport, and storage of neurotransmitters and neuromodulators; insertion of membrane channel proteins and receptors; and neuromodulatory effects. ①–⑫. These summarize the more rapid steps involved in immediate signaling at the synapse. These steps are described in the text, and are further discussed for different types of synapses in Chapter 8. Abbreviations: IP₃, inositol triphosphate; CAM II, Ca/calmodulin-dependent protein kinase II; DAG, diacylglycerol; PK, protein kinase; R, receptor; G, G protein; AC, adenylate cyclase.

Fig. 8 : mécanisme de la transmission chimique interneuronale (MDC, 2000, p.9)

¹⁵⁹ Machamer, Peter, Lindley Darden, and Carl Carver (2000), *Thinking About Mechanisms*, p. 3.

1) Continuité productive

La définition de MDC met en évidence une caractéristique importante du mécanisme : la production constante et continue de changements réguliers entre deux intervalles de temps¹⁶⁰. Autrement dit, un mécanisme exerce une transformation régulière et sans interruptions entre deux états donnés. Les conditions de possibilité de cette régularité sont triples : 1) les conditions externes du mécanisme sont inchangées 2) les conditions initiales sont similaires et 3) l'organisation interne est identique. En effet, si l'input des mécanismes est spécifique et que le mécanisme possède une structure stable, l'output sera régulier (toutes choses étant égales). Considérant ces trois conditions remplies par un système donné, la production de la relation input-output sera régulière. Nous verrons comment cette notion de continuité productive permet le phénomène épistémique de l'intelligibilité.

Remarquons que la définition classique de loi de la nature n'est pas utilisée pour saisir la régularité des mécanismes. Selon Hempel, une loi de la nature est universelle (sans exception), vraie, supporte les contrefactuels et permet la déduction dans l'explication (une condition supplémentaire, peu relevée dans la littérature, est celle de compatibilité (Feyerabend, 1975) : un énoncé universel sera considéré comme une loi s'il s'accorde logiquement avec la théorie acceptée)¹⁶¹. Une telle notion de loi n'est pas applicable en biologie puisque les régularités les plus fondamentales possèdent leurs exceptions¹⁶².

Cette remarque sur l'inapplicabilité de la théorie hempelienne de la loi servira aussi à MDC d'argument contre l'exigence de la loi causale directe dans la définition du mécanisme de Glennan¹⁶³. Nous pensons que le désaccord est simplement

¹⁶⁰ Machamer, Peter, Lindley Darden, and Carl Carver (2000), *Thinking About Mechanisms*, p. 3.

¹⁶¹ Hempel, Carl (1966) *Éléments d'épistémologie (Philosophy of Natural Science)*, p. 83-90.

¹⁶² Par exemple, la synthèse des protéines peut avoir lieu sans partie codante correspondante dans l'ADN à cause du phénomène de l'épissage. La loi de la nature de Hempel est même rejetée par le post-nagelien Schaffner (1993) pour une notion faisant abstraction du critère d'universalité.

¹⁶³ Les parties de tout mécanisme sont immédiates mécanistiquement et sont susceptibles d'être caractérisées par des lois d'action immédiate (chapitre 2, p. 81).

terminologique puisque Glennan n'utilise pas la notion de loi dans le sens de Hempel. En effet, nous avons vu que la notion de loi causale directe est plutôt une formalisation ou une généralisation¹⁶⁴ de l'interaction entre deux parties qui sont immédiates mécanistiquement. Or, MDC soutiennent, sous un vocabulaire différent, une idée identique. Ils reconnaissent l'importance de l'idée d'immédiacité mécanistique par l'entremise de la notion équivalente, que nous venons de décrire, de continuité. En effet, la notion de continuité stipule que les interactions entre les parties ne doivent pas être discontinues, c'est-à-dire disjointes, ce qui correspond à la notion l'immédiacité mécanistique de Glennan (p. 66-68). De plus, ils soutiennent aussi, nous le verrons dans la partie *Les explications causales-mécanistes*, que les formalisations entre les différentes parties d'un même mécanisme sont une composante importante de l'explication scientifique puisqu'elles peuvent servir à contextualiser un mécanisme. Nous pouvons donc conclure qu'il n'existe pas de désaccords entre Glennan et MDC quant aux lois, et donc au rôle des formalisations, et à la notion de continuité ou immédiacité mécanistique.

2) Entité et Activité

Au niveau ontologique, MDC pensent que l'ensemble des éléments invoqués en biologie moléculaire et en neurosciences tombe sous deux catégories, soit les entités et les activités. Les neurones, les membranes, les neurotransmetteurs et les ions sont des entités alors que la dépolarisation, la diffusion, le transport et la conduction sont des activités. Ces dernières sont responsables de la production des changements dont les entités sont les composantes. MDC adoptent donc une ontologie dualiste où la réalité est composée d'entités et d'activités. Plus précisément, la pertinence du dualisme ontologique provient du fait que deux entités peuvent être engagées dans

¹⁶⁴ La formalisation se distingue de la loi par sa portée : la formalisation est une expression algébrique d'une régularité naturelle d'un mécanisme particulier qui ne peut servir, à elle seule, d'explication par subsomption.

plusieurs activités différentes et une activité peut-être réalisée par plusieurs types d'entité.

Remarquons, une fois de plus, la divergence concernant la théorie ontologique de Glennan, qui est substantialiste, et MDC, qui optent pour un dualisme ontologique. Pour comprendre leurs positions, il faut brièvement rappeler le cadre originel du débat, les enjeux qui opposent *substantialistes* et les *processualistes* (*process ontologists*) concernant les entités fondamentales. Ce problème a été soulevé suite à une double tentative : comprendre les activités comme épiphénomènes de l'interaction entre les propriétés des entités et de l'autre côté, une priorité métaphysique envers les activités (processus) dont découle un traitement déflationniste des entités – ce ne serait que des abstractions commodes.

Les substantialistes privilégient les entités car ils pensent possible de reconstruire les activités à partir des seules entités. Plus simplement, en connaissant les entités et leur organisation, il est possible d'en inférer l'activité qui en résultera. Cartwright, par exemple, pense rendre compte des processus causaux (et pas uniquement d'association) à l'aide des capacités que possèdent les entités¹⁶⁵.

Les processualistes partagent l'opinion contraire¹⁶⁶. D'après Rescher (2000), le temps et le changement sont les catégories les plus importantes pour saisir la réalité. Conséquemment, les processus sont considérés l'élément ontique premier (option aussi choisie par Salmon pour résoudre les problèmes de l'atomisme phénoméniste de Hume). Cette orientation théorique permet un argument épistémologique : l'activité

¹⁶⁵ «It is not the laws that are fundamental, but rather the capacities. Nature selects the capacities that different factors shall have and sets bounds on how they can interplay. Whatever associations occur in nature arise as a consequence of the actions of these more fundamental capacities. In a sense, there are no laws of association at all. They are epiphenomena.» in Cartwright, N. (1989) *Nature's Capacities and Their Measurement*, p. 181.

¹⁶⁶ «As process philosophers see it, processes are basic and things derivative, because it takes a mental process (of separation) to extract 'things' from the blooming buzzing confusion of the world's physical processes. Traditional metaphysics sees processes (such as the rod's snapping under the strain when bent sufficiently) as the manifestation of dispositions (fragility), which must themselves be rooted in the stable properties of things. Process metaphysics involves an inversion of this perspective. It takes the line that the categorical properties of things are simply stable clusters of process-engendering dispositions. » Rescher, N (2000) *Process Philosophy*, p. 7.

est ce qui permet aux choses de nous être connues car l'activité permet d'engendrer des effets que nous pouvons détecter par nos sens. D'ailleurs, de multiples phénomènes sont étudiés en tant que simples processus, sans une analyse centrée sur leurs entités (par exemple, la fluctuation du champ magnétique terrestre, l'affaiblissement du champs gravitationnel solaire, ...). Bref, « One can only observe what things do – through their discernable effects – what there are, over and above this, is a matter of theory projected on this basis. »¹⁶⁷.

La stratégie œcuménique de MDC est de reconnaître une contrainte mutuelle entre entité et activité dans les limites de la biologie moléculaire et des neurosciences : les propriétés des entités limitent le nombre possible d'activité et une activité ne peut pas être déclenchée par n'importe quelle entité. La nature de ces interactions peut être très variable selon les entités et activités considérées. Une activité comme la coordination motrice peut être accomplie par de multiples macro-entités : une large partie des êtres vivants et un nombre important, et croissant, d'artefacts. À l'inverse, une activité comme la conscience réflexive n'est instanciée que par très peu d'organismes (les humaines et peut-être quelques grands mammifères) mais aucun artefact. Les limitations fixées par les entités sont aussi importantes. Une activité comme la vie n'est permise qu'en vertu d'une organisation physique précise : un sucre de désoxyribose, un groupe phosphate et de quatre bases. En neuroimagerie, certains scientifiques utilisent ce dualisme ontologique en supposant que des *entités* (isotope de l'oxygène) peuvent permettre de localiser des zones d'*activités* cérébrales (par le BOLD).

Malgré l'existence des contraintes induites par les entités, MDC accordent une attention particulière aux activités. Nous avons déjà mentionné que les activités sont les «producers of change»¹⁶⁸. Le terme «activité» ne renvoie pas à la notion de causalité conçue comme simple régularité, comme le soutenait les empiristes logiques

¹⁶⁷ Resher, N. (2000) *Process Philosophy*, p. 8.

¹⁶⁸ Machamer, Peter, Lindley Darden, and Carl Carver (2000), *Thinking About Mechanisms*, p. 4.

(suite à la critique humienne de la causalité). Ce sont les activités qui induisent, selon les propriétés des entités, les phénomènes de transformations régulières qui caractérisent les mécanismes. Comprendre ce que signifie la nature active d'un mécanisme peut s'effectuer grâce à un contraste. Une approche négative de la notion d'activité s'appliquerait à un système agrégatif. En effet, un agrégat est un exemple typique d'un système non-actif. L'activité peut être vue comme le phénomène qui nie l'intersubstituabilité, la constance qualitative, l'invariance après une reconstitution et l'absence de relation (simulatrice ou inhibitrice) entre les parties¹⁶⁹.

Ainsi, la continuité productive n'est pas interprétée comme une loi de la nature qui permet l'explication mais, à l'inverse, une caractéristique réclamant une explication ; explication fournie par une analyse mécaniste, c'est-à-dire l'étude empirique des relations entre les parties grâce à la compréhension du mode d'action des activités sur les entités et réciproquement.

Cette prise de position rappelle celle de Salmon (1984) concernant l'importance des processus (et de la notion d'interaction) dans l'analyse de la causalité : les régularités naturelles ne doivent pas servir à expliquer par subsomption mais plutôt être considérées comme des explanandum, des choses à expliquer (par le recours à des modèles de mécanismes). Toutefois, la méthode d'analyse de la causalité de MDC est radicalement différente de Salmon : en suivant Anscombe (1971), qui affirme que la notion de causalité ne possède de signification précise qu'autant qu'elle est remplacée par des termes plus spécifiques, MDC ne reconnaissent comme cause que les forces telles qu'elles sont décrites par les sciences. Les analyses philosophiques de la causalité, comme celles tentées par Salmon (1984), Mackie (1974) ou, plus récemment, par Glennan (1992, 1996) et Kistler (1999)) seront

¹⁶⁹ Les quatre critères qui font d'un système un système agrégatif : Wimsatt (1986), p. 262, 264, 266, 269.

abandonnées au profit de notions causales scientifiques particulières¹⁷⁰. Autrement dit, la stratégie consiste à utiliser les termes scientifiques spécifiques qui sont informatifs (conditions d'applications précises et support quantitatif) plutôt qu'un mot peu significatif tant il s'applique à une kyrielle de situations. Selon MDC, il existe quatre types de modes causaux en neurobiologie et en biologie moléculaire¹⁷¹ :

Activités	Exemples	Domaines d'investigation
Géométrico-mécanique	transport cellulaire	Chimie (chimie physique)
Electrochimique	action enzymatique	Biochimie
Énergétique	diffusion	Thermodynamique
Électromagnétique	système de navigation	Biophysique, Physique

Fig. 9 : typologie des types d'activités pour les neurosciences

Il existe d'autres types d'activités mais MDC se limitent aux forces jugées significatives, actuellement, pour l'étude au niveau moléculaire¹⁷². Les descriptions de mécanismes neurobiologiques ont donc une limite inférieure (*bottom out*) qui correspond, approximativement, à la biochimie (la limite inférieure de celle-ci étant approximativement la chimie physique¹⁷³).

¹⁷⁰ Le désintérêt face à une approche philosophique de la causalité est maintenu par Darden et Craver dans leurs écrits postérieurs à 2000 mais dans *Activities and Causation* (2002) Machamer revendique l'insuffisance d'une détermination scientifique de la causalité.

¹⁷¹ Machamer, Peter, Lindley Darden, and Carl Carver (2000), *Thinking About Mechanisms*, p. 14.

¹⁷² Toutefois, « The most important kinds of activities in molecular biology are geometrico-mechanical and electro-chemical activities. » in, Darden, Lindley (2002) *Strategies for Discovering Mechanisms: Schema Instantiation, Modular Subassembly, Forward/ Backward Chaining*, p. S356.

¹⁷³ Les ouvrages de référence en biochimie se limitent généralement aux principes de la thermodynamique (par exemple, Voet et Voet (1998)).

3) Organisation spatiotemporelle

Reconnaître qu'une organisation physique comme un mécanisme possède une conformation spatiotemporelle n'est pas une trivialité : la compréhension des propriétés matérielles est l'un des champs les plus vastes de l'entreprise scientifique humaine¹⁷⁴. D'autant plus que le paradigme naturaliste, et le néo-mécanisme doit sûrement être considéré comme une figure de proue naturaliste (c'est du moins la thèse générale de ce mémoire), a pour ambition d'expliquer des phénomènes aussi complexes que la cognition à partir des seules propriétés matérielles, ce qui fait reposer la viabilité du naturalisme sur l'étendue et la force explicative de nos conceptions sur ces propriétés matérielles¹⁷⁵.

MDC, et surtout Craver et Darden, partagent l'opinion que la connaissance philosophique ne doit pas se limiter à reconnaître simplement que la matière est organisée spatiotemporellement mais plutôt à fournir des stratégies de recherche à des fins de découvertes scientifiques. En cela ils suivent le projet philosophique de Wimsatt, Simon, Bechtel et Richardson, pour ne citer que les principaux, et se donnent pour mandat d'extraire de l'histoire des sciences des heuristiques liées aux contraintes spatiales et temporelles. Ainsi que nous l'avons montré pour Bechtel et Richardson dans le deuxième chapitre, leur projet est de renverser le préjugé philosophique selon lequel il n'existe pas de logique (au sens faible de rationalité) de la découverte. Concrètement, il s'agit de mettre à jour les contraintes empiriques qui existent entre les composantes des mécanismes (entités, activités) et les comportements de ces mécanismes.

¹⁷⁴ Croire le contraire peut-être imputé au phénomène de limite inférieure (*bottom out*) dans le cadre d'une philosophie autonomiste : on ne s'aventure qu'en des occasions trop peu nombreuses dans la compréhension des diverses propriétés de la matière.

¹⁷⁵ Rappelons la dépendance de l'importance du mécanisme face aux sciences naturelles : « The store of entities, activities, and modules out of which mechanisms are constructed expands and contracts with the addition and removal of established entities and activities over time. » in Craver, Carl. F. et Darden, Lindley (2001) *Discovering Mechanisms in Neurobiology: The Case of Spatial Memory*, p. 123.

Deux exemples, tirés de l'histoire de la physiologie du XIX^{ème}, permettent d'illustrer les importants progrès qui peuvent être accomplis en se souciant des propriétés temporelles des mécanismes. Il s'agit de montrer que le temps est une contrainte pertinente aux modèles mécanistes qui sont sensés expliquer les phénomènes naturels.

Johannes Müller (1801-1858), bien que reconnu comme un des pères fondateurs de la physiologie, a conservé une théorie des esprits animaux quant à la transmission nerveuse. Une fâcheuse conséquence de cette croyance fut qu'il crut que la transmission nerveuse, étant donné sa grande vitesse, ne pouvait être mesurée (la croyance en la spontanéité de la transmission nerveuse était aussi répandue, surtout en rapport aux facultés psychologiques – une des raisons pour laquelle Kant condamna toute étude empirique de la psychologie). Son célèbre élève Hermann Von Helmholtz réfuta cette idée par un raisonnement et un procédé expérimental simple. En examinant le différentiel temporel d'une contraction musculaire proximale et distale, il réussit à calculer la vitesse de l'influx nerveux : 30 m/s¹⁷⁶.

Une découverte fondamentale pour la neurobiologie fut la découverte des synapses par C. S. Sherrington (1906). Nous avons, à présent, adopté une vue analytique du système nerveux où celui-ci est composé d'entités interliées, les neurones. Or, au début de ce siècle, cette théorie, dénommée neuronique (Cajal), était opposée à la théorie réticulaire de Golgi (le système nerveux est continu). Les méthodes d'observation se trouvant dans un état d'avancement insuffisant pour une corroboration *de visu*, il n'a été possible de trancher définitivement la question qu'à la naissance de la microscopie électronique.

Toutefois, Sherrington a postulé l'existence de la discontinuité du système nerveux suite à une expérience en principe similaire à celle de Helmholtz. En connaissant la vitesse de la conduction nerveuse (200 pieds/s), et en connaissant la

¹⁷⁶ La réaction du père de Helmholtz face à ce résultat est largement reprise par beaucoup d'ouvrages, mais elle est trop irrésistible pour ne pas être retranscrite, encore une fois : « ... I could as little reconcile myself to your view as I could admit a star that had disappeared in Abraham's time should still visible. » in Churchland (1986), p. 24.

longueur d'un arc réflexe particulier (2 pieds), le temps de la transmission aurait dû être de 10ms en supposant l'arc continu. Toutefois, le temps réel de la transmission dans l'arc réflexe fut dix fois supérieur, soit 100ms. Sherrington en conclut qu'il devait exister entre les neurones sensoriels et les neurones moteurs (il s'agit d'un arc réflexe) une complexité structurale ralentissant la conduction de l'influx : ce fut le mécanisme hypothétique de la synapse, un espace interneuronal reliant les nerfs sensitifs et les nerfs moteurs. La discontinuité structurale des arcs réflexes était donc appuyée par des données temporelles. La conclusion importante pour nous est que le temps est une donnée informative sur les entités et les activités qui composent les mécanismes¹⁷⁷.

Plus pragmatiquement, Craver et Darden (2001) nous proposent de classer les activités selon quatre catégories temporelles : l'ordre, le taux, la durée et la fréquence. Nous ne décrirons que l'ordre qui est le type le plus important pour l'intérêt général du néo-mécanisme. En effet, l'ordre de succession des composants d'un mécanisme est étroitement lié à une conception linéaire (souvent formalisée par : $A \rightarrow B \rightarrow C$). Il s'agit bien évidemment d'une méconnaissance de la diversité des mécanismes : ils peuvent être circulaires (ce qui permet les mécanismes cycliques), en fourche, etc.¹⁷⁸. Ceci est d'une extrême importance, car il ne s'agit pas seulement de reconnaître la diversité pour elle-même mais en ce qu'elle constitue un outil de recherche : le mécanisme classique, qui ne postulait que des mécanismes linéaires, a développé des modèles biologiques qui se sont tous avérés faux¹⁷⁹. Comme nous l'avons montré pour la découverte de la vitesse de conduction nerveuse et les synapses, limiter les mécanismes à un type temporel particulier (linéaire) restreint arbitrairement la portée explicative et heuristique du mécanisme. Rappelons que la faillite du mécanisme cartésien est due à la pauvreté des activités reconnues par Descartes et, de la même

¹⁷⁷ *Ibid.*, p. 22-24.

¹⁷⁸ Voir les interactions fondamentales de Salmon au chapitre 1.

¹⁷⁹ Voir Descartes (1618-1637).

façon, reconnaître les mécanismes circulaires est capital pour la compréhension des organismes vivants en ce qu'ils sont les fondements du phénomène de l'homéostasie.

D'une façon tout aussi évidente, de multiples contraintes spatiales s'imposent aux mécanismes : la compartimentation, la localisation, la connectivité, la structure et l'orientation¹⁸⁰. Les éléments d'un mécanisme compartimenté (la transcription de l'ADN), s'exécutent dans un lieu ayant des limites physiques assez manifestes (le noyau). Un tel phénomène facilite le travail des scientifiques en limitant le nombre possible au nombre réel d'entités et d'activités (ainsi que leur organisation) qui constituent le mécanisme.

Lorsque les limites naturelles d'un mécanisme sont ambiguës, c'est-à-dire non compartimentées, il peut être tenté de localiser le mécanisme (les méthodes expérimentales seront abordées à la prochaine partie). Ce procédé permet de bénéficier des avantages de la compartimentation. Nous avons déjà constaté la pertinence du travail de Bechtel et Richardson (1993) à ce titre¹⁸¹. Mais, cette heuristique étant faillible, la localisation n'assure pas de frontière aussi certaine que la compartimentation. En fait, les nombreux cas dans l'histoire de la biologie et de la biochimie que nous avons rapportés semblent indiquer qu'il faut être patient dans l'élaboration de la localisation des mécanismes, pire, de multiples tentatives sont généralement nécessaires – dans le cas où la localisation est finalement une réussite.

La connectivité est une contrainte physique importante et peut être grandement utile, comme le montre le cas de l'étude de l'hippocampe. Ce dernier était considéré comme étant un composant important du système olfactif dans les années 1940-50. Quelques évidences d'inhibition, sur des modèles animaux, montrent que l'ablation de l'hippocampe entraîne des conséquences négatives importantes. Ce rôle fonctionnel attribué à l'hippocampe sera définitivement rejeté lorsqu'il sera

¹⁸⁰ Darden, Lindley (2002) *Strategies for Discovering Mechanisms: Schema Instantiation, Modular Subassembly, Forward/ Backward Chaining*, p. 124-127.

¹⁸¹ Une intéressante question empirique pourrait être adressée : 1) peut-on évaluer l'utilité pour un chercheur de connaître cette heuristique et 2) l'aide-t-il réellement ?

découvert qu'aucune connexion n'existait entre celui-ci et le bulbe olfactif (Craver, 2004a).

La structure et l'orientation sont aussi des propriétés hautement décisives. Une belle illustration de l'importance des facteurs structuraux et d'orientation est le cas du premier élément de la vie elle-même : la création spontanée de la membrane grâce aux acides gras (phosphoglycérolipides) en milieu aqueux.

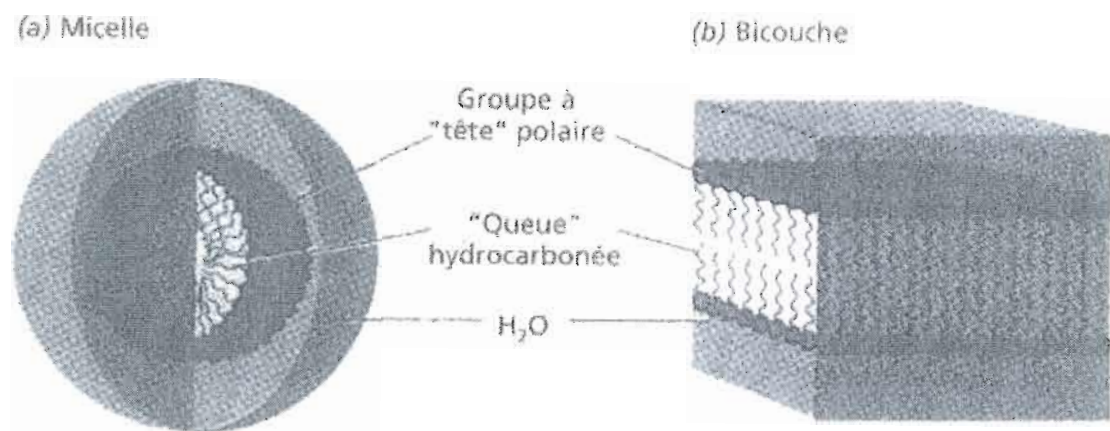


Fig. 10 : micelles et bicouches (Voet, Donald et Voet, Judith (1998) *Biochimie*, p.33)

Les acides gras sont composés d'une « tête » hydrophile et d'une « queue » hydrophobe. Ces propriétés mènent les acides gras à former des micelles ou des bicouches, c'est-à-dire des arrangements organisés de façon conformes à la charge électronique : les queues des acides gras seront attirées les unes vers les autres en raison de leur affinité alors que les têtes seront attirées par l'eau. Le résultat est la création d'un milieu interne et d'une perméabilité sélective qui est la base du transport passif entre les milieux intracellulaire et extracellulaire. Ainsi, en vertu de l'orientation des entités (molécules d'eau et acides gras) suite à la nature des activités (liaison covalente : attirer et repousser) se forme une structure organisée.

En somme, la caractérisation ontologique d'un mécanisme s'appuie sur une description spatiotemporelle de ses éléments constitutifs, les entités et les activités.

Loin de reconnaître les caractéristiques matérielles des constituants comme étant d'importance secondaire, l'approche de MDC consiste plutôt à montrer que les propriétés spatiotemporelles de l'organisation des mécanismes sont centrales à la recherche scientifique. L'agencement des entités et des activités est ce qui détermine la transformation régulière d'un état *a* à un état *b*. Autrement dit, expliquer ce qu'un mécanisme fait, c'est montrer comment l'arrangement spatial et temporel des entités et des activités permet certains comportements productifs.

II. Les caractéristiques épistémiques du mécanisme

Nous avons décrit ce qu'est un mécanisme. Il nous faut, à présent, montrer comment nous nous représentons un mécanisme car expliquer ou décrire est nécessairement une activité cognitive. Pour un agent, expliquer se fait par l'intermédiaire d'un schéma mécaniste. C'est une représentation abstraite d'un mécanisme. Un schéma est donc une simplification mais une simplification qui permet toutefois de comprendre la continuité productive caractéristique des mécanismes. Le degré d'abstraction varie selon la nature du mécanisme étudié¹⁸², l'état de la recherche et le but visé par le chercheur. Remarquons que nous pourrions formuler les mêmes critiques à la notion d'abstraction de MDC qu'à la notion de similarité de Glennan : les néo-mécanistes optent pour une approche purement pragmatique au sujet de l'évaluation des modèles par rapport aux mécanismes eux-mêmes – sans toutefois montrer pourquoi il faudrait renoncer à tout formalisme.

Nous expliciterons donc les principales caractéristiques épistémiques (schémas et sketches mécanistes, niveau et hiérarchie mécaniste) des explications mécanistes afin de préciser une théorie causale-mécaniste de l'explication (intégration des explications étiologiques, contextuelles et constitutives) et nous exposerons,

¹⁸² Vouloir comprendre mécanistiquement la conscience est certainement plus difficile que de comprendre mécanistiquement les mouvements squelettiques (ce qui a été accompli par l'école iatomécanique dès le XVII^{ème} et XVIII^{ème}).

brièvement et principalement, les méthodes expérimentales qui mènent à leur production.

1) Les schémas mécanistes

La complexité des systèmes biologiques exige des outils pour se représenter les multiples avenues par lesquelles les mécanismes sont liés entre eux dans la production de phénomènes aussi élaborés que, par exemple, la mémoire. Comme nous l'avons suggéré plus haut, une fonction importante du schéma mécaniste est l'intelligibilité : un schéma permet de mettre en évidence les connexions structurales entre les parties et leur organisation temporelle. Une explication mécaniste est donc l'élaboration d'un schéma mécaniste qui montre comment les entités et activités sont organisées pour produire de façon régulière un phénomène particulier. Les schémas possèdent donc les caractéristiques généralement attribuées aux théories : « they are discovered, evaluated, and revised in cycles as science proceeds. They are used to describe, predict, and explain phenomena, to design experiments, and to interpret experimental results. »¹⁸³.

La première remarque importante au sujet des schémas mécanistes concerne leurs limites épistémiques. Ils ont des conditions de départ et des conditions finales. Le rôle de l'explication mécaniste est de montrer comment la constitution du mécanisme permet d'obtenir l'organisation de l'état initial à l'organisation de l'état final. Déterminer les limites temporelles des schémas mécanistes est une opération délicate en ce qu'elle est une perspective particulière sur les mécanismes étudiés. En ce sens, elle est très contextuelle et il existe plusieurs découpes possibles d'un même mécanisme. Autrement dit, il peut exister une pluralité de schémas d'un seul

¹⁸³ Machamer, Peter, Lindley Darden, and Carl Carver (2000), *Thinking About Mechanisms*, p. 16-17.

mécanisme¹⁸⁴. Le contexte d'analyse d'un mécanisme est le fruit de l'appartenance disciplinaire du chercheur, de ses orientations théoriques, de ses moyens techniques, de ses préjugés, ... bref, de facteurs psychologiques, techniques, théoriques et institutionnels¹⁸⁵. Résumons en rappelant que les schémas mécanistes sont pragmatiquement établis dans d'étroites limites – avec l'inévitable flou qu'entraîne une réponse aussi pragmatiste.

La deuxième remarque importante est à propos de la nature représentative des schémas mécanistes. Alors que les empiristes logiques conçoivent l'explication comme une inférence déductive entre des propositions (chapitre 1), les explications mécanistes sont mieux représentées par des diagrammes. De nombreux mécanismes sont complexes : ils sont eux-mêmes composés de mécanismes, font partis de mécanismes plus généraux et ils sont souvent non-linéaires (cycliques). Or, les représentations picturales ont des ressources symboliques plus nombreuses et plus efficaces que les simples mots (et les propositions) pour représenter une telle complexité. Pensons à la commodité de représenter les contraintes spatiales (la compartimentation, la localisation, la connectivité, la structure et l'orientation) par de simples reproductions graphiques simplifiées des entités. Une telle entreprise serait un véritable labeur sous une forme propositionnelle : imaginez un manuel de neurosciences présenté uniquement sous une forme propositionnelle !

D'ailleurs, Bechtel et Abrahamsen (2005c) présentent un argument en faveur du mécanisme basé sur une théorie de l'esprit et de la représentation : le mode d'explication mécaniste posséderait des affinités avec notre propre cognition. En effet, Bechtel et Abrahamsen (2005a) sont plus explicites que MDC sur les artefacts représentationnels à la disposition des chercheurs. Notamment, ils soulignent les

¹⁸⁴ Nous reviendrons plus précisément sur ce qu'est une explication mécaniste dans la section *Les explications causales-mécanistes*, p. 111.

¹⁸⁵ Pour des exemples de la pluralité possible des schémas d'un même mécanisme, voir p. 71, 72 et 118.

limites des diagrammes¹⁸⁶ : ils sont statiques. S'appuyant sur un modèle simulationniste de la cognition, ils abordent le travail de Heagarty (1992) et Barsalou (1998) pour décrire comment nous faisons des inférences ressemblant à des animations mentales qui présentent l'incontestable avantage d'être dynamiques (surtout pour les représentations temporelles)¹⁸⁷. Penser serait donc une récupération, dans la mémoire à long terme, de catégories par des proxytypes (concrètement, les proxytypes peuvent être des représentations multi-modales, de simples représentations visuels ou des représentations de mots¹⁸⁸). Réfléchir, imaginer, ne sont plus alors que des réactualisations de configurations originales de l'état cérébral recherché. Au sens littéral, le simulationnisme soutient une thèse représentationnaliste selon laquelle penser et la re-présentation, par la médiation de la mémoire de travail et les proxytypes, des événements du monde.

Bref, selon Bechtel et Abrahamsen (2005c), les théories simulationnistes apportent un support cognitif aux caractéristiques épistémiques du néo-mécanisme telles que présentées dans les revues et les ouvrages de référence scientifiques. Ainsi, bien que MDC n'abordent pas explicitement ce genre de légitimation cognitive, par la supposée affinité du mécanisme avec notre cognition effective, ils n'en tirent pas moins la leçon : Darden, en particulier, construit des modèles animés de mécanismes génétiques¹⁸⁹.

Les schémas mécanistes sont donc des artefacts représentationnels internes (animation mentale ou simulation) ou externes (diagrammes, animations, propositions) qui sont fixés pragmatiquement en vue d'un objectif de recherche et

¹⁸⁶ Nombreux sont ceux qui affirment que les schémas picturaux ont de multiples et puissants avantages sur les schémas strictement linguistiques : Notamment, Hegarty, Simon, Langley, Bradshaw, Zytrow, Larkin, Stenning, Lemon, Bechtel et Abrahamsen (pour les références, voir Bechtel, W. and Abrahamsen, A. (2005c) *Explanation: A Mechanistic Alternative*).

¹⁸⁷ «Frequently diagrams provide a preferred representation of a mechanism. Inference involves a determination of how a mechanism behaves, and this is typically not achieved via logical inference but by simulating the activity of a mechanism, either by animating a diagram or by creating mental, computational, or scale model simulation. » in Bechtel and Abrahamsen (2005c) *Explanation: A Mechanistic Alternative*, p. 11).

¹⁸⁸ Prinz, J. J. (2001) *Furnishing the Mind*, p. 149.

¹⁸⁹ <http://www.philosophy.umd.edu/Faculty/LDarden/Research/index.html>

dont la fonction est de rendre intelligible un phénomène en l'expliquant (en fournissant une organisation d'entité et d'activité ayant une continuité productive).

2) Les sketches mécanistes

Les schémas sont donc un support clef pour l'explication. Cependant, le schéma ne peut pas toujours rendre compte de la continuité productive d'un phénomène mais seulement véhiculer quelques informations plus ou moins éparses et dont quelques composants sont non-analysés (il sont alors dénommés des boîtes noires). MDC réservent le terme de « sketch mécaniste » pour une telle situation. Schéma et sketch mécaniste se distinguent donc par le critère de continuité productive. Les sketches sont utiles aux chercheurs en ce qu'ils indiquent les insuffisances qui pourraient faire d'un sketch un schéma. De plus, le rassemblement, même incomplet, des données sur un sujet d'étude peut servir à contraindre, même minimalement, les hypothèses de schéma mécaniste.

3) Niveau et hiérarchie mécaniste

Une des caractéristiques principales des explications mécanistes est qu'elles impliquent plusieurs niveaux organisés hiérarchiquement (chapitre 1). Illustrons cet aspect à l'aide d'un thème important de la littérature sur la philosophie des neurosciences, tant au niveau de l'explication que du problème de la réduction (Bickle 2002, Schouten, 2005), la mémoire. Darden/Craver (2001) et Craver (2002) s'appliquent surtout à l'explication intégrée de la mémoire spatiale grâce à plusieurs niveaux d'analyse. Voici une représentation graphique :

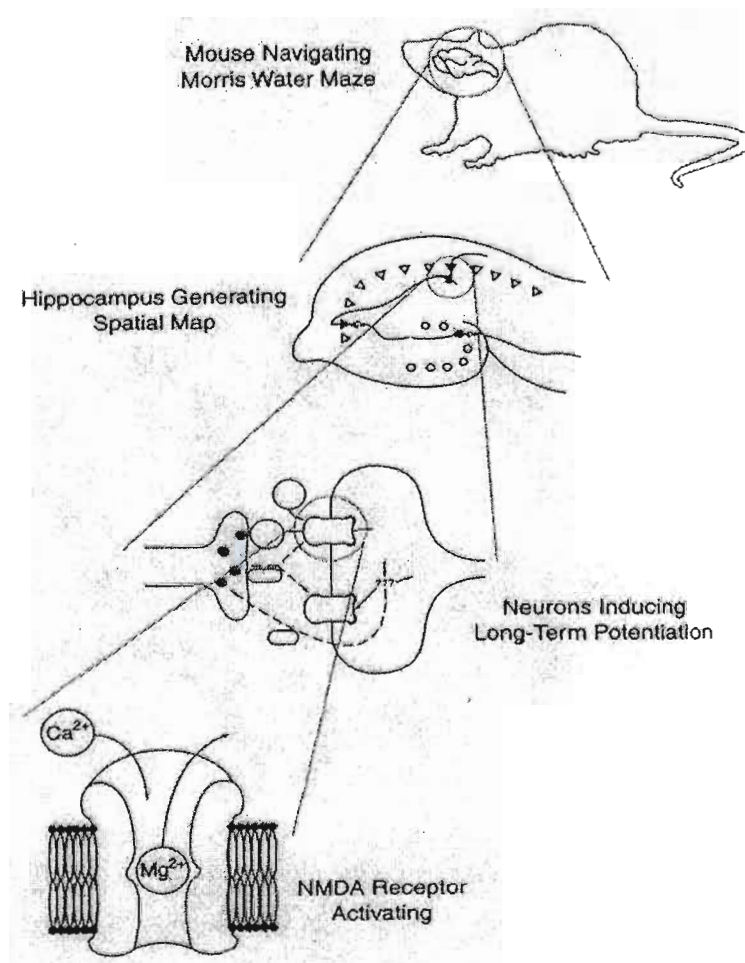


Fig. 11 : représentation multi-niveaux hiérarchisée de la mémoire spatiale¹⁹⁰

Le graphique ci-dessus indique le processus de décomposition mécanistique et correspond à un sketch mécaniste puisqu'il ne répond pas au critère de continuité productive, de nombreuses étapes étant inconnues et certains composants encore insuffisamment compris (notamment le mécanisme qui représente la limite inférieure, celui de la potentialisation à long terme, PLT¹⁹¹). Néanmoins, il donne une certaine idée de la structure dont dépend la capacité d'un rat à s'orienter dans l'espace, la

¹⁹⁰ Craver (2002), p. S90.

¹⁹¹ Pour avoir une idée de l'état des débats entourant la PLT (LTP en anglais) : ce mécanisme a été dénommé *Long Term Controversy* ou *Long Term Tar Pit* (Malinow 1998), in Craver (2001), p. S86.

navigation, en fournissant une analyse constitutive de cette capacité. Brièvement, l'orientation spatiale est due à la maîtrise de « carte spatiale » située dans l'hippocampe. L'investigation se poursuit : certaines hypothèses ont été émises sur la compréhension des cartes spatiales en termes d'entités et d'activités propres à l'hippocampe, principalement, son organisation cytologique. Les travaux de E. Kandel rendent possible un autre niveau d'analyse, encore plus poussé, celui du renforcement interneuronal par le mécanisme de PLT.

Cette hypothèse mécaniste sur la capacité d'orientation spatiale du rat implique donc quatre niveaux distincts mais imbriqués de façon hiérarchique : l'orientation est due aux cartes, particulièrement à leur organisation cytologique dans l'hippocampe, elle-même redevable, en partie, au mécanisme de PLT. Nous ne nous intéresserons pas aux détails ni à l'évaluation de la valeur d'une telle hypothèse. Il nous importe surtout de dégager le fait que les explications sont des articulations compositionnelles de niveaux.

Remarquons que Craver (2001) se distingue de Wimsatt (et de Bechtel (1986)) en ce qu'il fait interpréter « niveau » de façon épistémique et non pas ontologique. Pour lui, les niveaux mécanistes sont des perspectives descriptives¹⁹² et non pas des niveaux se référant à un niveau de la nature. D'ailleurs il transparaît que les niveaux en question (comportemental, anatomique, cytologique, moléculaire) sont établis en fonction des sciences actuelles (psychologie/éthologie, neuroanatomie, neurobiologie et biologie moléculaire) et des ressources techniques dont nous disposons (protocoles comportementaux, microscopie/imagerie, *BatchClan*/Électrophysiologie/Intervention pharmacologique, *Knock out*/*Polymerase Chain Reaction*). En principe, si l'on possédait les moyens de décrire au niveau quantique le comportement de la PLT, il pourrait alors être rajouté à l'analyse multi-niveaux présentée. De façon plus réaliste, on pourrait envisager d'intégrer des

¹⁹² Craver (2001b), p. 65-66.

contraintes de types environnementales car celles-ci affectent manifestement le phénomène de la mémoire.

Bref, la nature pragmatique de l'analyse mécaniste s'applique à la notion de niveau et, dans cette perspective, il est utile de distinguer trois types de niveau selon le type d'analyse dont il est fait usage : 1) niveau agrégatif, 2) niveau fonctionnel et 3) niveau mécaniste. Les trois sens de niveau s'appliquent à caractériser des relations méréologiques, c'est-à-dire des relations tout-partie. Les deux premières s'appliquent, respectivement, aux rapports entre les entités (volume/masse) et aux rapports entre les activités (conçues de façon abstraite). Seul le niveau mécaniste s'applique à une relation méréologique qui se fonde sur le dualisme ontologique de MDC (qui s'applique aux entités et aux activités).

Les agrégats, comme nous l'avons exposé dans le deuxième chapitre, sont des systèmes où le comportement du tout n'est pas modifié par la réorganisation des parties. La notion de niveau agrégatif fait donc l'économie de l'étude des activités. Pour l'étude des relations spatiales lors d'une décomposition, cette notion de niveau agrégatif peut être utile. Toutefois, cet avantage est bien évidemment très limité dans la mesure où les systèmes biologiques sont des systèmes intégrés.

Comprendre ce qu'est un niveau fonctionnel s'établit dans le cadre d'une analyse¹⁹³. Le résultat d'un fractionnement d'une tâche en un ensemble de sous-tâches coordonnées procure une description abstraite du fonctionnement d'un système. Le rôle de chacune des sous-tâches permet d'obtenir la capacité analysée du système en question. Dans cette perspective, la notion de niveau fonctionnel est une relation entre des rôles différents¹⁹⁴ : un tel niveau d'abstraction implique que le niveau fonctionnel ne s'attarde pas aux entités réelles du système réel (fonctionnalisme). Remarquons cependant que l'analyse fonctionnelle est plus privilégiée en

¹⁹³ « ... the analytical strategy proceeds by analyzing a disposition d of a into a number of other disposition $d_1 \dots d_n$ had by a or components of a such that programmed manifestation of the d_i results in or amounts to a manifestation of d », in Cummins, Robert (1984) *Functional Analysis*, p. 402.

¹⁹⁴ Craver (2002) *Interlevel Experiments and Multilevel Mechanisms in the Neuroscience of Memory*, p. S88.

psychologie, les neuroscientifiques s'occupant plutôt de l'instanciation des descriptions abstraites du fonctionnement d'un système (au moyen des techniques expérimentales) mais, comme le souligne judicieusement Schouten (2005)¹⁹⁵, un travail d'instanciation doit être ouvert à tout changement au niveau de l'analyse fonctionnelle. Ultimement, la solution paraissant idéale serait celle où s'exercerait des contraintes mutuelles entre les divers domaines.

La notion de niveau qui nous préoccupe est celle de niveau mécaniste. Précisément, un niveau est, la plupart du temps, présenté comme une perspective sur un mécanisme (schéma ou sketch). Cette mise en perspective peut provenir d'une part, d'un niveau supérieur, d'autre part, d'un niveau inférieur. Comme nous l'avons déjà remarqué, les explications neuroscientifiques sont typiquement distribuées sur plusieurs niveaux mécanistiques ; ceci permet d'illustrer la hiérarchie des mécanismes et, à différents degrés de précision, le fonctionnement interne de certaines capacités. Le niveau supérieur et le niveau inférieur ont le même référent (le phénomène étudié), mais ils le décrivent de façon à illustrer comment un certain mécanisme produit un certain comportement. Les niveaux mécanistes permettent donc une intégration de l'explanandum dans un niveau supérieur (comportemental, par exemple) et dans un niveau inférieur (en procédant à l'étude de ses entités et de ses activités) en identifiant 1) le contexte dans lequel s'inscrit le mécanisme étudié et 2) les éléments constitutifs du mécanisme étudié. Nous allons voir à présent comment situer l'explanandum dans une hiérarchie de niveaux compositionnels permet d'expliquer mécanistiquement.

4) Les explications causales-mécanistes

Nous savons à présent que nous nous représentons les mécanismes par des schémas ou des sketches mécanistes qui sont paradigmatiquement distribués sur

¹⁹⁵ Le point soulevé par Schouten concerne la réduction de la mémoire au mécanisme de la PLT : il affirme que les partisans d'une telle réduction se basent sur des modèles psychologiques dépassés.

plusieurs niveaux, constituant ainsi une hiérarchie mécaniste. Nous allons maintenant nous attarder à montrer comment est conçue une explication neuroscientifique dans ce contexte.

Les types étiologiques et constitutifs d'explication nous sont déjà familiers suite à la section sur la *Mechanical Philosophy* de Salmon. Craver (2001b, 2002, 2004a) fait valoir la nécessité d'un autre genre d'explication, les explications contextuelles. Ce rajout se situe toutefois en droite ligne avec l'intuition de Salmon sur ce qu'est fondamentalement une explication : « Scientific explanation, according to the ontic conception, consists in exhibiting the phenomena-to-be-explained as occupying their places in the patterns and regularities which structure the world. »¹⁹⁶.

L'explication contextuelle de Craver est amenée expressément pour parvenir à mieux situer un mécanisme dans son réseau causal. En effet, nous avons vu que les mécanismes sont polyfonctionnels ; un même mécanisme peut être utilisé dans plusieurs explications différentes de phénomènes différents. Cela est particulièrement important pour les mécanismes largement répandus parmi les êtres vivants puisqu'ils sont employés dans diverses utilisations : rappelons-nous le mécanisme la dépolarisation. Un autre parallèle pourrait être fait avec la notion d'exaptation de Gould. Ce terme désigne le « recyclage » d'un organe afin d'exercer une nouvelle fonction¹⁹⁷. Un processus similaire est à l'œuvre pour les mécanismes.

La remarque qui s'impose alors est que situer un mécanisme exige que l'on rapporte le contexte dans lequel il exerce son rôle causal. Ceci est d'autant plus important que nous avons souligné le caractère hiérarchique et multi-niveaux des explications neuroscientifiques. D'ailleurs, une autre motivation de l'introduction de l'explication contextuelle dans les explications causales-mécanistes est qu'elle permet l'intégration des fonctions dans une analyse mécaniste, sans pour autant être réduites

¹⁹⁶ Salmon (1984) *Scientific Explanation and the Causal Structure of the World*, p. 239.

¹⁹⁷ L'exemple classique est l'exaptation des ailes en tant que régulateur thermique aux ailes dont la fonction est de permettre le vol.

à une explication étiologique télescopée¹⁹⁸. En suivant Cummins (1984), Craver affirme que les fonctions ont un pouvoir explicatif, ce sont des explications contextuelles, car elles informent l'agent du rôle causal que joue le mécanisme dans le système où il se trouve. Autrement dit, « they are descriptions of the activity of some item explicitly in terms of how it is organized into the workings of a higher (+1) level mechanism. »¹⁹⁹.

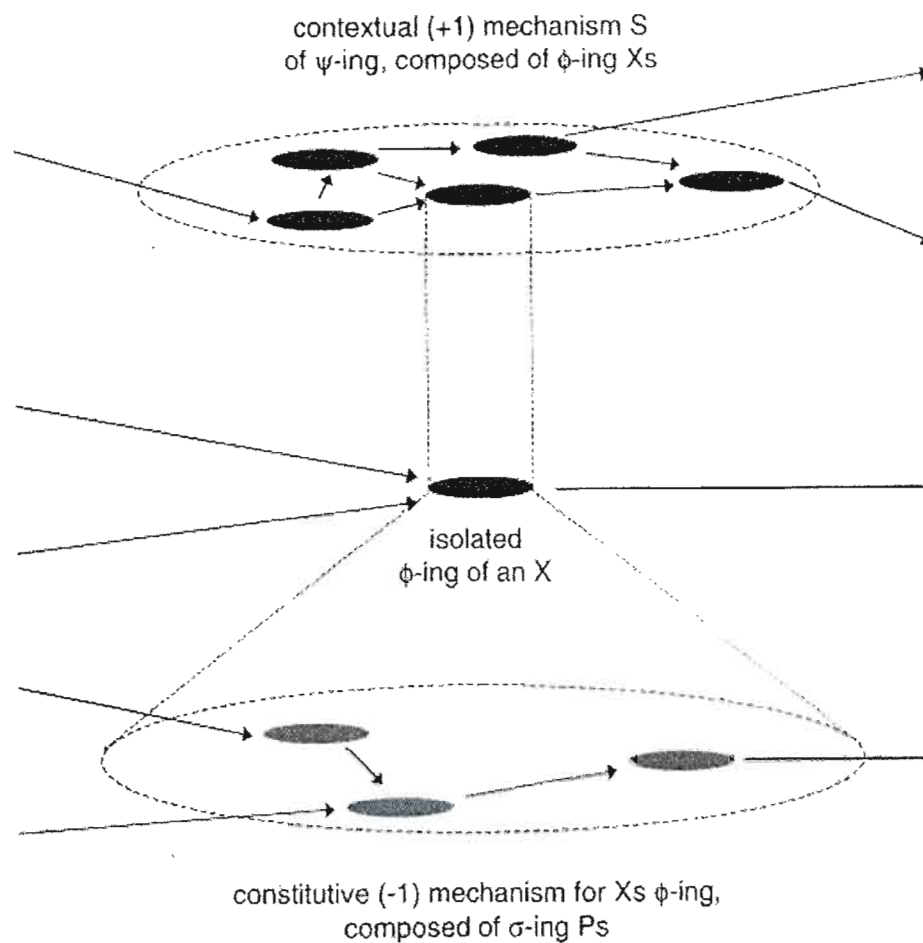


Fig. 12 : l'explication mécaniste (Craver, 2001)

¹⁹⁸ Thèse défendue par Hempel (1965-1966), Nagel (1961), Wright (1976) et Salmon (1984).

¹⁹⁹ Craver (2004a) *Functions and Mechanisms in Contemporary Neuroscience*, p. 18.

En bref, une explication causale-mécaniste est l'intégration de trois types d'explications, les explications étiologiques, constitutives et contextuelles. Ensemble, elles permettent de tracer un portrait de la hiérarchie mécaniste. Les explications étiologiques, qui sont limitées aux causes proximales en neurosciences²⁰⁰, sont représentées par les conditions initiales du niveau analysé et où les conditions finales représentent le phénomène à expliquer. L'explication constitutive est sensée illustrer comment le fonctionnement interne du mécanisme est responsable de la transformation qui a eu lieu entre les deux limites temporelles établies. Finalement, l'explication contextuelle doit permettre à l'agent de situer le rôle du mécanisme étudié, ce qui est souvent fait en donnant la participation de ce mécanisme dans un mécanisme plus général (la dépolarisation dans la PLT, la dépolarisation dans la simple transmission nerveuse, ...). Craver utilise les expressions imagées de *backward looking*, *downward looking* et *upward looking*, pour les explications étiologiques, constitutives et contextuelles²⁰¹, pour renforcer l'idée qu'une explication situe, réellement, un événement dans un réseau causal. L'explication étiologique permettrait de situer l'explanandum en déterminant ses causes antérieures (*backward looking*), l'explication constitutive permettrait de préciser mécanistiquement l'explanandum (*downward looking* ou *inward looking*) et, finalement, l'explication contextuelle permettrait de fixer le contexte dans le lequel l'explanandum opère (*upward looking*).

Le résultat recherché est l'intégration des entités et des activités à différents niveaux²⁰². Intégrer un mécanisme exige que l'on soit en mesure de fournir une explication contextuelle et une explication constitutive. Supposons que le mécanisme

²⁰⁰ Les neurosciences n'ont pas comme mandat de se prononcer, à l'aide d'explications distales, sur le pourquoi de la présence d'un mécanisme. À vrai dire, ce n'est pas le mandat de grand monde : une telle entreprise est tellement spéculative, et nous ne voyons pas comment elle pourrait l'être moins, qui ne vaut certainement pas la peine d'y consacrer du temps et des ressources. Nous n'y avons tout simplement aucune prise empirique.

²⁰¹ Craver (2004a)(soumis) *Functions and Mechanisms in Contemporary Neuroscience*, p. 24.

²⁰² Craver (2002) *Interlevel Experiments and Multilevel Mechanisms in the Neuroscience of Memory*, p. S91.

x est étudié au niveau (0). Intégrer x signifie connaître au niveau (+1) son rôle causal et au niveau (-1), l'organisation de ses entités et activités qui permettent de produire la capacité de ce mécanisme x . L'intégration des mécanismes permet à un sketch de devenir un schéma mécaniste, dans la mesure où l'on peut tracer le chemin ou la voie causale (*pathway*) des niveaux (-1) au niveau (+1) sans fossé dans l'explication, c'est-à-dire en suivant sans interruption le comportement des entités et activités.

5) Les stratégies et méthodes expérimentales

Comment l'établissement d'une intégration de telles hiérarchies mécanistes est-il possible de fait ? Par les stratégies et méthodes expérimentales multi-niveaux. Le principe d'une expérience multi-niveaux est simple : il s'agit 1) d'avoir établi un schéma mécaniste, 2) d'y appliquer des techniques d'interventions et 3) d'enregistrer les influences provoquées par les interventions (par comparaison avec des données références). L'on peut distinguer deux types d'expérimentations : les expérimentations ascendantes (*bottom up*) et descendantes (*top down*). Les expérimentations ascendantes sont nommées ainsi car les techniques d'intervention ont lieu dans des niveaux mécanistes inférieurs alors que les techniques de détection sont placées sur les voies spécifiées par l'hypothèse de recherche dans les niveaux supérieurs. À l'inverse, les techniques descendantes affectent les niveaux supérieurs et sont suivies et détectées sur les voies inférieures. Le graphique de la page suivante illustre les deux types d'expérimentation pour un seul niveau.

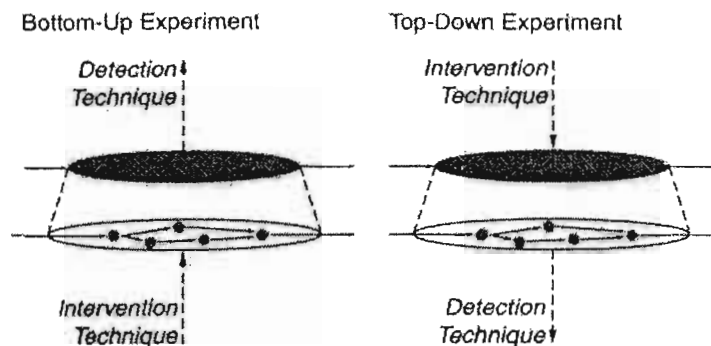


Fig. 13 : expérimentations ascendantes (*bottom up*) et descendantes (*top down*)²⁰³.

Craver relève trois stratégies d'expérimentation : 1) les stratégies d'activation, 2) les stratégies d'interférence et 3) les stratégies adjuvantes (*additive*)²⁰⁴. La première a pour finalité de localiser les mécanismes sous-jacents d'une capacité particulière. La stratégie d'activation est donc *top-down*. Bien souvent, l'activation est déclenchée et contrôlée par des tâches comportementales standardisées et à l'aide de techniques de détection (*Single Cell Recording* ou imagerie mentale), l'on tente de discerner les parties qui sont liées spécifiquement à l'activité en cours. Un des arguments pour la thèse localisationniste de la mémoire dans l'hippocampe est établi sur la base d'une stratégie d'activation : les enregistrements provenant d'électrodes sur 70 à 150 cellules pyramidales de l'hippocampe permettent de prédire le chemin emprunté par un rat dans son labyrinthe (Craver, 2002)²⁰⁵.

Les deux stratégies suivantes sont *bottom up*. L'interférence est une intervention qui modifie négativement le comportement d'une entité ou d'une activité

²⁰³ Craver (2002) *Interlevel Experiments and Multilevel Mechanisms in the Neuroscience of Memory*, p. S92.

²⁰⁴ Bechtel et Richardson (1993) catégorisent les types d'intervention en études inhibitrices et excitatrices (p. 20). Elles correspondent aux stratégies d'interférence et adjuvante. La stratégie d'activation est traitée dans Bechtel (à paraître) *The Epistemology of Evidence in Cognitive Neuroscience* et Bechtel et Stufflebeam (1996).

²⁰⁵ La référence originale est Wilson, M.A. et Bruce McNauhghton (1993) *Dynamics of the Hippocampal Ensemble Code for Space*, Science 261:1055-1058, in Craver (2002).

afin de tenter de discerner quels en sont les effets au niveau supérieur²⁰⁶. L'expression « modification négative » renvoie à un ensemble de pratiques inhibitrices : de la simple diminution à l'élimination complète. La destruction complète des entités et activités ne s'avère informative que depuis assez récemment : des techniques génétiques peuvent supprimer le gène NMDAR1²⁰⁷, qui code en partie pour les récepteurs à glutamate, dans l'hippocampe (les souris traitées sont appelées « knockout mice »). Ce gène particulier est ciblé car il est impliqué dans le mécanisme de la PLT. Les récepteurs à glutamate jouent un rôle essentiel dans le renforcement interneuronal ; or, il s'est avéré que ces souris ont eu de graves difficultés de navigation²⁰⁸.

La stratégie adjuvante est antagoniste à la stratégie de l'interférence en ce qu'elle est une excitation des activités ou une amplification des entités. Un second gène code en partie pour les récepteurs NMDA, le NMDA2, qui est lui-même composé de sous-unités, appelons-les, pour des raisons de commodité, *x* et *y*. L'expression de *x* et *y* est régulée, notamment, par l'âge de la souris : les jeunes souris expriment plus *x*, les vieilles, plus *y*. Les chercheurs ont artificiellement augmenté l'expression de *x* et des tests expérimentaux ont montré que les souris transgéniques apprennent plus rapidement et plus efficacement (elles retiennent l'information plus durablement)²⁰⁹.

Nous avons très brièvement illustré les trois stratégies expérimentales utilisées par les neuroscientifiques. Bien évidemment elles possèdent leurs inconvénients et ils ne sont pas moindres. Rapidement, les *top-down* activatrices recèlent une grave difficulté : le problème des causes communes. Ce qui est détecté (enregistrement de

²⁰⁶ Bickle s'est servi de telles études pour tenter de montrer que la suppression d'un type récepteur n'affecte, selon une batterie de tests expérimentaux, qu'une certaine forme de conscience. L'étude qu'il a présentée servait à illustrer que la neurobiologie moléculaire était un niveau pertinent pour l'étude de la conscience.

²⁰⁷ NMDA signifie N-Methyl D-Aspartate.

²⁰⁸ Les références originales sont McHugh et al. 1996, Rotenberg et al. 1996, Tsien et al. 1996 et Tsien et al. 1996), in Craver (2002).

²⁰⁹ Référence originale Tang et al. (1999) *Genetic Enhancement of Learning and Memory in Mice*, Nature, 401: 63-69, in Craver (2002).

l'activation d'un neurone ou le BOLD) peut n'être qu'un épiphénomène par rapport à la véritable cause. Pour cette raison ce genre d'étude n'est bien souvent que préliminaire à des méthodes plus fiables, celles qui offrent un meilleur contrôle des voies biochimiques ou moléculaires. L'avantage de cette méthode est principalement heuristique en indiquant des covariations entre une tâche comportementale et une région plus ou moins ciblée. La stratégie d'interférence a le problème évident d'infliger des changements structuraux (chirurgie)²¹⁰ ou biochimique (pharmacologie) qui peuvent biaiser les résultats. Notons que ce genre de problèmes tend à s'amoindrir avec l'affinement de techniques d'intervention. Des remarques similaires peuvent être faites pour la stratégie adjuvante.

Une conclusion importante sur les méthodes expérimentales multi-niveaux est qu'elles sont toujours utilisées de manière croisée. Les neuroscientifiques ont la possibilité de confronter leurs hypothèses de schémas mécanistes sur les phénomènes étudiés grâce aux trois méthodes, comblant ainsi les défauts inhérents à chaque méthode. De plus, le fait qu'une entité donnée soit manipulée par des techniques et s'inscrit dans des cadres théoriques différents, lui confère une certaine robustesse.

Nous avons donc présenté les principales caractéristiques épistémiques (schémas et sketches mécanistes, niveau et hiérarchie mécaniste) des mécanismes afin de préciser une théorie causale-mécaniste de l'explication (intégration des explications étiologiques, contextuelles et constitutives) et nous avons exposé, brièvement et en principe, les méthodes expérimentales qui permettent de les créer (les stratégies d'activation, les stratégies d'interférence et les stratégies adjuvantes).

Voilà comment se présente la seconde théorie exhaustive du néo-mécanisme. Les auteurs ont divisé leurs thèses en analyse métaphysique et analyse épistémologique du mécanisme. Ontologiquement, un mécanisme est un arrangement spatiotemporel d'entités et d'activités ou les activités sont les producteurs

²¹⁰ Pensons au principe de l'Unité du système nerveux (anti-localisationniste) de Flourens et à ses arguments expérimentaux...

responsables du changement productif qui caractérise le dynamisme des mécanismes. Au plan épistémique, en tant qu'agent cognitif, nous n'avons accès qu'à des représentations de mécanismes, les schémas et les sketchs. Les schémas mécanistes ont un pouvoir explicatif dans la mesure où ils montrent comment l'organisation interne est responsable de la continuité productive des mécanismes. Aussi, les schémas sont typiquement multi-niveaux en ce qu'ils sont distribués sur différentes perspectives d'un phénomène. L'entreprise de MDC doit être vue comme entreprise descriptive du travail effectif des neuroscientifiques et des biologistes. Leurs spéculations métaphysiques sur la nature ontologique du mécanisme peuvent être interprétée comme la mise en évidence des présupposés de la pratique scientifiques et non pas comme représentatives d'une prise de position des scientifiques. En ce sens, seule l'analyse épistémologique est descriptive alors que l'analyse métaphysique est simplement l'image que doit avoir le monde s'il doit être en accord avec nos croyances et nos actions.

Conclusion

Ce mémoire nous aura donc permis de comprendre quatre aspects importants de l'une des plus complètes théories de l'explication scientifiques actuellement disponible.

D'abord, le premier chapitre nous a permis de comprendre comment le néo-mécanisme a pu se construire suite à la philosophie des empiristes logiques. De façon générale, cette transition s'est avérée être une opposition : le néo-mécanisme se caractérise par un tournant réaliste et naturaliste par rapport à l'approche linguistico-formelle des empiristes logiques. Le rôle de Salmon se doit d'être une fois de plus souligné : il fut le premier à dénoncer la phobie anti-métaphysique des empiristes logiques en proposant une théorie ontique – et non pas seulement épistémologique – des explications causales scientifiques. Ce revirement réaliste a profondément modifié les exigences liées à l'explication : l'objectif du scientifique n'est plus de subsumer des événements à expliquer sous des lois universelles (liaison formelle) mais de circonscrire l'influence de certains processus causaux sur une structure matérielle particulière (liaison informelle). À ce titre, le contraste entre les explications causales-mécanistes telles que développées par MDC (sur le modèle original de Salmon) et les explications selon le modèle D-N met à jour la trop grande simplicité de ce dernier et rappelle une des principales leçons véhiculées par les néo-mécanistes : l'explication est une élucidation ardue des relations empiriques (et donc informelles) entre l'explanans et l'explanandum.

Le second tournant par rapport à l'empirisme logique, mené par Simon et Wimsatt, se caractérise par une attitude naturaliste et pragmatique face à la recherche scientifique. Parce que le projet du naturalisme pragmatiste s'intéresse à la compréhension des processus cognitifs responsables de la découverte en science, ce projet renverse l'approche de l'empirisme logique (rejet du contexte de découverte, concentration sur le contexte de justification) en intégrant la psychologie et l'intelligence artificielle dans l'épistémologie et la philosophie des sciences. Leur

apport ne se limitera pas la naturalisation de l'épistémologie puisque Simon et Wimsatt se sont penchés, parallèlement, sur l'ontologie sous-jacente aux sciences actuelles, la théorie des systèmes complexes. Nous avons vu comment ces auteurs soutiennent le mécanisme en défendant, par la rationalité limitée, la théorie ontologique selon laquelle les systèmes complexes sont arborescents et donc quasi-décomposables. Cette stratégie argumentative est courante dans le mouvement mécaniste : si la thèse ontologique de la quasi-décomposabilité est vraie, alors le mécanisme en tant que modèle de l'explication scientifique possède des assises épistémologiques réelles et solides.

Remarquons qu'une des motivations importantes pour l'adoption de cette nouvelle théorie ontologique (plutôt qu'un physicalisme tel que proposé par Salmon) ressortit probablement à la nécessité d'intégrer les entités reconnues par les sciences biologiques et non pas seulement les entités de la physique. Ainsi que nous l'avons souligné, la légitimité scientifique n'est plus l'apanage des sciences physiques et il ne s'agit pas d'un simple conflit disciplinaire mais bien plus de la reconnaissance de la complexité des systèmes vivants. Rappelons aussi que les néo-mécanistes refusent simplement, en adoptant l'ontologie des systèmes complexes, l'éliminativisme concomitant au physicalisme réductionniste.

D'ailleurs, cette reconnaissance de l'importance des sciences biologiques se manifeste à un degré aigu chez les néo-mécanistes, à tel point que leur philosophie des sciences est certainement une des premières à être développée complètement indépendamment des sciences physiques. Pensons à Glennan qui va même jusqu'à marginaliser la microphysique – en l'excluant du domaine des phénomènes mécanistiquement explicables – au point de l'isoler du reste de l'activité scientifique. Pensons aussi à Bechtel, Richardson et MDC pour lesquels la physique n'est que d'une utilité extrêmement limitée pour comprendre comment sont menées

effectivement les explications scientifiques²¹¹. Le néo-mécanisme, dans sa forte dépendance à l'ontologie scientifique, nous rappelle à quel point ce que l'on considère comme légitime scientifiquement en philosophie s'est étendu largement au-delà de la physique pour englober la biochimie, la biologie et la neuropsychologie.

Cet état de fait nous rappelle le premier avantage de l'approche mécaniste : elle constitue une norme d'intelligibilité de la nature. Or, force est de constater que pour les néo-mécanistes, les modèles de l'explication scientifique calqués sur les sciences physiques (le modèle laplacien ou le modèle D-N) sont inapplicables à la sphère des sciences qui s'occupent des êtres vivants – et, parmi ceux-ci, l'humain. Nous pouvons donc conclure que le néo-mécanisme procure une norme d'intelligibilité de la nature animée, ce qui était inaccessible aux modèles fondés sur la physique. En bref, le néo-mécanisme nous procure une norme d'intelligibilité incommensurablement plus développée pour les systèmes vivants (dont l'esprit humain) que ne l'ont fait les modèles de l'explication basés sur la physique.

Le deuxième aspect important couvert par ce mémoire a trait au processus de la découverte effective des mécanismes par l'activité scientifique. À l'encontre des empiristes logiques, il s'avère qu'analyser ce processus de la découverte scientifique n'échappe pas à toute logique, comme l'ont soutenu ces philosophes²¹². La réponse de Bechtel et Richardson, qui s'inscrit dans l'esprit de Simon et la rationalité limitée, consiste en une famille d'heuristiques représentant les étapes typiques de la recherche scientifique : 1) isoler le système de son milieu naturel et tenter de localiser un locus de contrôle pour la fonction étudiée au sien de ce système, 2) quand l'étape 1) est réussie, il faudra ensuite tenter de décomposer ce locus et de localiser des composants particuliers responsables de la production de cette fonction (si la décomposition et la localisation échoue, le programme mécanistes s'arrête ici), 3) un premier résultat de cette opération de décomposition et de localisation est l'indépendance de premier

²¹¹ Bechtel (2005) s'occupe de sciences physiques à une seule occasion, lorsqu'il évalue les caractéristiques et les limites de la neuroimagerie – qui repose sur la microphysique.

²¹² Voir la note 9 au chapitre 1.

ordre (le comportement du système est du à un seul composant, celui identifié), 4) un second résultat de cette opération de décomposition et de localisation est l'interaction de premier ordre : le comportement du système est du à plusieurs composants, il faut donc comprendre les interactions entre les composants du système, 4) un troisième résultat de cette opération de décomposition et de localisation est la reconceptualisation : ce résultat advient lorsque les interactions entre les composants sont trop complexes pour être comprises – la reconceptualisation de l'analyse à un niveau inférieur devient alors une solution possible.

Ces étapes cruciales de la recherche scientifique ne réfute pas seulement la position selon laquelle rien d'intéressant ne peut être extrait du processus du contexte de découverte mais procure un guide heuristique inestimable. De plus, les heuristiques du mécanisme démystifie la découverte en la réintroduisant dans une perspective de résolution de problème : découvrir un mécanisme est une entreprise éminemment complexe mais n'en demeure pas moins le long résultat d'une activité de résolution de problèmes par un agent cognitif contraint par de multiples facteurs.

Le troisième aspect important couvert par ce travail est la clarification d'une notion laissée, jusqu'à maintenant, à la simple intuition. Nous nous sommes rendus compte que d'importants débats ontologiques offrent deux perspectives concurrentes sur les éléments à inclure dans la notion de mécanisme. La première théorie exhaustive du néo-mécanisme s'appuie sur un monisme ontologique qui ne reconnaît que des entités qui interagissent entre elles alors que la seconde théorie exhaustive du mécanisme propose de conceptualiser le mécanisme à partir d'un dualisme ontologique (il existe des entités et des activités). Glennan et MDC reconnaissent que le fait important devant être considéré est la nature productive des mécanismes ; Glennan pense pouvoir en rendre compte à partir des seules entités ainsi que des relations entre elles alors que MDC croient plutôt que ce sont les activités qui servent à l'expliquer. Pour notre part, nous concluons que le desideratum important d'un point de vue mécaniste, c'est que chaque version 1) reconnaisse l'importance de la

nature productive des mécanismes et 2) en propose une explication vraisemblable.

Dans le même ordre d'idée, nous avons conclu que l'apparente dissension entre Glennan et MDC au sujet de l'importance des lois dans la formulation de ce qu'est un mécanisme n'était qu'un faux débat dans la mesure où Glennan 1) entend loi au sens de généralisation empirique et non de conditionnelle universelle et 2) se sert de cette notion pour spécifier la nécessité que nos représentation des mécanismes soient continus (il faut minimiser les discontinuités entre les entités et activités et, idéalement les abolir), ce qu'il exprime par l'expression d'immédiacité mécanistique. Ainsi, l'important pour Glennan et MDC est l'exigence de fournir une histoire causale continue – et non pas discontinue. Autrement dit, l'idéal régulateur de toute explication est l'obtention d'un schéma mécaniste plutôt qu'un sketch mécaniste pour l'explanandum.

Nous concluons ce mémoire en rappelant comment le néo-mécanisme permet une certaine unification de nos connaissances, à l'instar des formes passées de mécanisme. En effet, les mécanistes ont, de tout temps, argumenté en faveur de leur propre approche en vantant la convergence entre l'ontologie et la façon dont nous l'expliquons (et la façon dont nous devons l'expliquer). Descartes soutenait qu'expliquer un phénomène revient à comprendre comment les éléments qui le constitue s'entrechoquent : toutes les variations comportementales (des choses) devaient être expliquées par des variations de structures des éléments constitutants. Si l'explication procède ainsi c'est parce que le monde est ainsi : il est composé d'unité discrète qui s'entrechoque. L'ontologie dicte l'épistémologie. Actuellement, Simon Wimsatt, Bechtel et Richardson soutiennent la viabilité et la pertinence du mécanisme par une théorie ontologique. Si l'on peut expliquer un comportement ou un phénomène en le décomposant pour comprendre son articulation interne, c'est justement parce qu'il est ainsi *fait* ; le monde lui-même possède une structure arborescente (argument des horlogers), il est structuré en systèmes intégrés hiérarchiquement. Si le monde était plutôt un tout fortement intégré, au point de

soutenir un holisme fort, alors il nous serait impossible d'arriver à une quelconque isolation ou découverte « locale ».

Ainsi, la thèse de l'unification est d'abord une thèse ontologique. Toute entité ou propriété doit être conçue dans un unique cadre ontologique (monisme). Cependant, comme nous l'avons montré, l'organisation naissant entre différents systèmes peut mener à de nouvelles propriétés. L'exemple de la navigation chez le rat est manifeste : une certaine organisation cytologique permet de former des cartes spatiales qui – en interaction avec l'environnement – permettent à leur tour le phénomène de la navigation. Une approche ontologique moniste mais hiérarchisée possède une affinité évidente avec la théorie tripartite de l'explication mécaniste (explications étiologique, contextuelle et constitutive). En effet, le néo-mécanisme est l'explication intégrative par excellence : l'approche néo-mécaniste permet de comprendre l'articulation entre les niveaux puisqu'elle affirme la nécessité 1) de contextualiser l'explanandum à un niveau supérieur (+1) et 2) de décomposer l'explanandum à un niveau inférieur (-1). L'unification de nos différents domaines de connaissance est donc due 1) à une théorie ontologique moniste (mais non-réductionniste) et 2) une théorie de l'explication qui intègre les niveaux en articulant leurs relations par des schémas ou *sketchs* mécanismes.

Concluons sur un enjeu central à ce mémoire, la portée du néo-mécanisme et son importance comme doctrine de l'explication scientifique. Nous avons amplement montré que le couplage entre l'ontologie systémique et les explications néo-mécanistes possède une grande capacité d'intégration (thème de l'unification). De plus, la structure même du mémoire a été construite en fonction de montrer l'importante complétude et la cohérence de cette perspective philosophique : il existe deux théories exhaustives du néo-mécanisme, c'est-à-dire que les néo-mécanistes déploient beaucoup d'efforts pour couvrir l'ensemble des questions philosophiques concernant les explications scientifiques. D'ailleurs, force est de constater que la philosophie contemporaine des sciences ravive d'intéressantes questions ontologiques

et métaphysiques tout en s'intéressant de très près aux détails de l'investigation scientifique, tant au niveau méthodologique (les techniques d'expérimentations) que psychologique (processus de la recherche et de la découverte). Après avoir subi les contrecoups relativistes des réactions anti-empirisme logique, il semble que le néo-mécanisme exemplifie enfin une théorie de l'explication cohérente et sensible à la plupart des aspects des questions philosophiques concernant les sciences – mais surtout, le monde lui-même, incluant nos esprits.

La portée du néo-mécanisme n'aura donc jamais été aussi grande que de nos jours puisqu'elle inclut la plupart de nos disciplines naturelles et s'étend progressivement aux sciences humaines. La psychologie, contra Descartes, propose de multiples approches mécanistes mais de façon plus surprenante, de nombreux auteurs (Elster (1989) ; Sperber (1996) ; Hedström et Swedberg (1998) ; Bunge (2003, 2004) appliquent le néo-mécanisme aux sciences sociales, particulièrement la sociologie.

Bibliographie

- Anscombe, Gertrude, E., M. (1971) *Causality and Determination*, in *Metaphysics and the Philosophy of Mind*, The Collected Philosophical Papers of G. E. M. Anscombe, v. 2. Minneapolis: University of Minnesota Press, 133-147.
- Barberousse, Anouk, Kistler, Max et Ludwig, Pascal (2000) *La philosophie des sciences aux XX^{ème} siècle*, Flammarion.
- Beauregard, M., Lévesque, J. & Bourgouin, P. (2001) *Neural correlates of the conscious self-regulation of emotion*, Journal of Neuroscience, 21: RC165 (1-6).
- Bechtel, William et Robert C. Richardson (1993) *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*, Princeton University Press.
- Bechtel, William, P. Mandik, J. Mundale et R. S. Stufflebeam, ed. (2001) *Philosophy and the Neurosciences*, Blackwell.
- Bechtel, William. (2001) *The compatibility of complex systems and reduction: A case analysis of memory research*, Minds and Machines, 11, 483-502.
- Bechtel, William. (2002) *Decomposing the brain: A long term pursuit*, Brain and Mind, 3, 229-242.
- Bechtel, William. et Abrahamsen, A. (2005a)(à paraître) *Mechanistic explanation and the nature-nature controversy*, Bulletin d'Histoire et d'Epistémologie des Sciences de la Vie.
- Bechtel, William. et Abrahamsen, A. (2005b)(sous presse) *Explanation: A Mechanistic Alternative*, Studies in History of Philosophy of the Biological and Biomedical Sciences.
- Bechtel, William (2005)(à paraître) *The Epistemology of Evidence in Cognitive Neuroscience* in R. Skipper Jr., C. Allen, R. A. Ankeny, C. F. Craver, L. Darden, G. Mikkelsen, and R. Richardson (eds.), *Philosophy and the Life Sciences: A Reader*. Cambridge, MA: MIT Press.

- Bernard, Claude (1865) *Introduction à l'étude de la médecine expérimentale*, rééd. 1984, Champs/Flammarion.
- Bickle, John (2003) *Philosophy of Neuroscience: A Ruthlessly Reductive Account*. Dordrecht: Kluwer Academic Publishers.
- Boas, Marie (1956) *The Establishment of the Mechanical Philosophy*, University of Massachusetts.
- Bogen, James (2001) *What We Talk About When We Talk About Causality*, University of Pittsburgh Philosophy of Science Archive, <http://philsci-archive.pitt.edu>.
- Bogen, James (2004)(soumis) *The Regularity/Generalization Question in Mechanistic Philosophy*, Studies in History and Philosophy of Biological and Biomedical Sciences.
- Bunge (2003) *Emergence and convergence*, University of Toronto Press.
- Bunge, Mario (2004) *How Does It Work ? The Search for Explanatory Mechanisms*, Philosophy of the Social Sciences, Vol. 34 No. 2, 182-210.
- Campaner, R. I. (2000) *Wesley Salmon, Causality and Explanation*, Erkenntnis, vol. 52, no. 1, pp. 121-125(5).
- Cartwright, Nancy (1989) *Nature's Capacities and Their Measurement*, Oxford: Oxford University Press.
- Churchland, Patricia, S. (1986) *Neurophilosophy*, The MIT Press.
- Craver, Carl, F. (2001a) *Structures of Scientific Theories*, pour Machamer, P. and Silberstein, M. éd. Blackwell Guide to the Philosophy of Science, Oxford: Blackwell.
- Craver, Carl, F. (2001b) *Role Functions, Mechanisms and Hierarchy*, Philosophy of Science, 68: 31-55.
- Craver, Carl. F. et Darden, Lindley (2001) *Discovering Mechanisms in Neurobiology*:

- The Case of Spatial Memory*, in P.K. Machamer, R. Grush, and P. McLaughlin (eds.), *Theory and Method in Neuroscience*. Pittsburgh, PA: University of Pittsburgh Press, 112-137.
- Craver, Carl (2002) *Interlevel Experiments and Multilevel Mechanisms in the Neuroscience of Memory*, *Philosophy of Science Supplemental* 69: S83-97.
- Craver, Carl et Bechtel, William (2004) *Explaining Top-Down Causation (Away)*, Draft Version.
- Craver, Carl, F. (2004a)(soumis) *Functions and Mechanisms in Contemporary Neuroscience*, pour L. Faucher éd., *Reader in the philosophy of the neurosciences*.
- Craver, Carl, F. (2004b)(soumis) *Beyond Reduction: Mechanisms, Multifield Integration and the Unity of Neuroscience*, *Studies in History and Philosophy of Biological*.
- Cummins, Robert (1984) *Functional Analysis*, Sober, Elliott (ed.) (1984) *Conceptual Issues in Evolutionary Biology: An Anthology*, 1st. ed. Cambridge, MA: MIT Press, pp. 477-508.
- Darden, Lindley (1992) *Strategies for Anomaly Resolution*, in R. Giere (ed.), *Cognitive Models of Science*, *Minnesota Studies in the Philosophy of Science*, Vol. 15. Minneapolis: University of Minnesota Press, p. 251-273.
- Darden, Lindley et Michael Cook (1994) *Reasoning Strategies in Molecular Biology: Abstractions, Scans and Anomalies*, in D. Hull, M. Forbes, and R. M. Burian (eds.), *PSA 1994*, v. 2. East Lansing, Michigan: Philosophy of Science Association, pp. 179-191.
- Darden, Lindley (1995) *Exemplars, Abstractions, and Anomalies: Representations and Theory Change in Mendelian and Molecular Genetics*, in James G. Lennox and Gereon Wolters (eds.), *Concepts, Theories, and Rationality in the Biological Sciences*. Pittsburgh, PA: University of Pittsburgh Press, p. 137-158.
- Darden, Lindley (1998) *Anomaly-Driven Theory Redesign: Computational Philosophy of Science Experiments*, in T.W. Bynum and J.H. Moor, *The Digital Phoenix: How Computers are Changing Philosophy*, New York: Blackwell Publishers, p. 62-78.

- Darden, Lindley (2002) *Strategies for Discovering Mechanisms: Schema Instantiation, Modular Subassembly, Forward/ Backward Chaining*, Philosophy of Science (Supplement PSA 2000 Part II) 69: S354-S365.
- Davies, Paul (1996, 1988) *Les forces de la nature*, trad. A. Bouquet, Champs Flammarion.
- Descartes, René (1618-1637) *Œuvres philosophiques* (I), *Traité de l'homme*, éd. F. Alquié, Classiques Garnier.
- Diderot, Denis (2002) *Le Rêve de d'Alembert*, GF Flammarion.
- Dijksterhuis, E. J. (1961) *The Mechanization of the World Picture*, trad. C. Dikshoorn, Oxford: Clarendon Press.
- Duchesneau, François (1998) *Les modèles du vivant de Descartes à Leibniz*, Vrin.
- Elster, John (1989) *Nuts and Bolts for the Social Sciences*, Cambridge University Press.
- Finger, Stanley (1994) *Origins of Neuroscience*, Oxford University Press.
- Feyerabend, Paul (1975, 1979) *Contre la méthode*, Seuil.
- Frank, Philipp et Shorr, Philip (1937) *The Mechanical versus the Mathematical Conception of Nature*, Philosophy of Science, Vol. 4, No 1, 41-74.
- Giere, Ronald, N. (1992) *Cognitive Models of Science*, Minnesota Studies in the Philosophy of Science, Vol. XV.
- Giere, Ronald, N. (2000a) *Cognitive Approaches to Science*, in Newton-Smith, W. H. (2000) *A Companion to the Philosophy of Science*, Blackwell Companions to Philosophy.
- Giere, Ronald, N. (2000b) *Naturalism*, in Newton-Smith, W. H. (2000) *A Companion to the Philosophy of Science*, Blackwell Companions to Philosophy.
- Glennan, Stuart, S. (1992) *Mechanisms, Models, and Causation*, Ph.D. Dissertation, Chicago, University of Chicago.

- Glennan, Stuart, S. (1996) *Mechanisms and The Nature of Causation*, *Erkenntnis* 44: 49-71.
- Glennan, Stuart, S. (1997) *Capacities, Universality, and Singularity*, *Philosophy of Science* 64:605-626.
- Glennan, Stuart S. (2002) *Rethinking Mechanistic Explanation*, *Philosophy of Science* (Supplement) 69: S342--S353.
- Hacking, Ian (1989) *Concevoir et expérimenter*, Christian Bourgois Éditeur, trad. B. Ducrest.
- Hall, Thomas S. (1969) *History of General Physiology*, The University of Chicago Press.
- Hedström, P. et Swedberg, R. (1998) *Social Mechanism, An Analytic Approach to Social Theory*, Cambridge University Press.
- Hempel, Carl, G. et Oppenheim, P. (1948) *Studies in the Logic of Explanantion*, *Philosophy of Science* 15: 135-175.
- Hempel, Carl (1965) *Aspects of scientific explanation and other essays in the philosophy of science*, New York : Free press.
- Hempel, Carl (1966) *Éléments d'épistémologie (Philosophy of Natural Science)*, trad. B. Saint-Sernin, Armand Colin.
- Hume, David (1748) *Enquête sur l'entendement humain*, trad. P. Baranger et P. Saltel, G-F.
- Humphreys, Paul, W. (1989) *Scientific Explanation: The Causes, Some of the Causes, and Nothing But the Causes*, in P. Kitcher et W. C. Salmon (ed.), *Minnesota Studies in the Philosophy of Science*, Vol. XIII.
- Jacob, Pierre (1996) *De Vienne à Cambridge*, Gallimard.
- Kistler, Max (1999) *Causalité et lois de la nature*, Vrin.

- Kitcher, Philip (1985) *Two Approaches to Explanation*, The Journal of Philosophy, Volume LXXXII.
- Kitcher, P. (1989) *Explanatory Unification and the Causal Structure of the World*, in P. Kitcher et W. C. Salmon (ed.), Minnesota Studies in the Philosophy of Science, Vol. XIII.
- Lafontaine, Céline (2004) *L'empire cybernétique*, Seuil.
- La Mettrie, Julien, O. (2000) *L'homme machine*, Mille et une nuits.
- Lenoir, Timothy (1982) *The Strategy of Life, Teleology and Mechanics in Nineteenth Century German Biology*, D. Reidel Publishing Company.
- Lévesque, J., Joannette, Y., Paquette, V., Mensour, B., Beaudoin, G., Leroux, J.-M., Bourgouin, P. & Beauregard, M. (2003) *Neural circuitry underlying voluntary self-regulation of sadness*, Biological Psychiatry, 53: 502-510.
- Machamer, Peter, Darden, Lindley, et Carver, Carl (2000) *Thinking About Mechanisms*, Philosophy of Science 67: 1-25.
- Machamer, Peter, K., Grush, Rick et McLaughlin, Peter (2001) *Theory and method in the neurosciences*, Pittsburgh : University of Pittsburgh Press.
- Machamer, P. et Silberstein, M. (2002) *Philosophy of Science*, Blackwell Philosophy Guides.
- Machamer, Peter (2002) *Activities and Causation*, PSA.
- Machamer, Peter et Sullivan, Jacqueline (2004) *Leveling Reductionism*, Draft version, philsci-archive.pitt.edu/archive/00000400/00/Leveling_reduction,8.24.doc.
- Mackie, John, L. (1974) *The Cement of the Universe: A Study of Causation*. Oxford: Oxford University Press.
- Mayntz, R. (2003) *Mechanisms in the Analysis of Social Macro-Phenomena*, Philosophy of the Social Sciences, Vol. 34, No. 2, June, 237-259.

- Nagel, Ernst (1950) *Mechanistic Explanation and Organismic Biology*, Philosophy and Phenomenological Research, Volume XI: 327-338.
- Nagel, Ernst (1961) *The Structure of Science*, New York : Harcourt, Brace and World.
- Newton-Smith, W. H. (2000) *A Companion to the Philosophy of Science*, Blackwell Companions to Philosophy.
- Papineau, David (1989) *Pure, Mixed, and Spurious Probabilities and Their Significance for a Reductionist Theory of Causation*, in in P. Kitcher et W. C. Salmon (ed.), *Minnesota Studies in the Philosophy of Science*, Vol. XIII.
- Popper, Karl, R. (1959) *The Logic of Scientific Discovery*, New York : Harper & Row.
- Prinz, Jesse, J. (2001) *Furnishing the Mind*, MIT Press.
- Reichenbach, Hans (1951) *The rise of Scientific Philosophy*, University of California Press.
- Reichenbach, Hans (1966, 1938) *Experience and Prediction*, University of Chicago Press.
- Rescher, Nicholas (2000) *Process Philosophy*, University of Pittsburgh Press.
- Salmon, Wesley C. (1984) *Scientific Explanation and the Causal Structure of the World*, Princeton University Press.
- Salmon, Wesley, C. (1989) *Four Decades of Scientific Explanation*, University of Minnesota Press.
- Salmon, Wesley, C. (1997) *Causality and explanation: A Reply to Two Critiques*, Philosophy of Science 64: 461-477.
- Salmon, Wesley, C. (1998) *Causality and explanation*, Oxford University Press.
- Sawyer, Keith, R. (2003) *The Mechanisms of Emergence*, Philosophy of the Social

Sciences, Vol. 34, No. 2, June, 260-282.

Schaffner, Kenneth (1984) *Reduction in Biology: Prospects and Problems*, *Scientific Explanation*, Sober, Elliott (ed.) (1984), *Conceptual Issues in Evolutionary Biology: An Anthology*, 1st. ed. Cambridge, MA: MIT Press.

Schaffner, Kenneth, F. (1993) *Discovery and Explanation in Biology and Medicine*, University of Chicago.

Schaffner, Kenneth (2004) *The Cheshire Cat Problem and a Return to Root*, Draft Version.

Schouten, Maurice, K. D. (2005)(à paraître) *The engram and beyond: The elimination of consolidation*, in Faucher, L., Poirier, P., & E. Ennan (Eds.). *Des Neurones à la Philosophie: Neurophilosophie et Philosophie des Neurosciences*, De Boeck.

Searle, John (2001) *Rationality in Action*, MIT Press.

Simon, Herbert, A. (1996) *Les sciences de l'artificiel*, trad. Le Moigne, Gallimard.

Sperber, Dan (1996) *La contagion des idées*, Odile Jacob.

Steel, Daniel (2004) *Empirical Analysis and Causal Pluralism; or, What in the world is Causal Structure ?*, présentation à l'APA, <http://www.msu.edu/user/steel/causality.pdf>.

Sterelny, Kim et Griffiths, Paul, E. (1999) *Sex and Death, An introduction to Philosophy of Biology*, The University of Chicago Press.

Stich, Stephen (1993) *Naturalizing Epistemology: Quine, Simon and the Prospects for Pragmatism*, in Hookway, Christopher et Peterson, Donald (1993) *Philosophy and Cognitive Science*, Royal Institute of Philosophy Supplement: 34, Cambridge University Press.

Tabery, J. (2004a) *Activities and Interactions*, *Philosophy of Science* 71: 1-15.

Tabery, J. (2004b) *Synthesizing Activities and Interactions in the Concept of a Mechanism*, Forthcoming in *Philosophy of Science*.

- Thagard, Paul (1998) *Explaining Disease: Causes, Correlations, and Mechanisms*, *Minds and Machines* 8: 61-78.
- Thagard, Paul (2003) *Pathways to Biomedical Discovery*, *PSA*, 70, p. 235-254.
- Van Fraassen, Bas, C. (1985) *Salmon on Explanation*, *The Journal of Philosophy*, Volume LXXXII.
- Von Eckardt, Barbara et Poland, Jeffrey, S. (2005)(à paraître) *Mechanism and Explanation in Cognitive Neuroscience*, *Philosophy of science*.
- Wagner, Pierre, dir. (2002) *Les philosophes et la science*, Gallimard.
- Westfall, Richard, S. (1977) *The Construction of Modern Science*, *Mechanisms and Mechanics*, Cambridge University Press.
- Wilson, Jessica (2004) *Causality*, *Routledge Encyclopedia of Philosophy of Science* (Draft of 6-17-04).
- Wimsatt, William (1972) *Complexity and Organization*, in Kenneth F. Schaffner and Robert S. Cohen (eds.), *PSA 1972, Proceedings of the Philosophy of Science Association*. Dordrecht: Reidel, pp. 67-86.
- Wimsatt, William (1976) *Reductive Explanation: A Functional Account*, in A. C. Michalos, C. A. Hooker, G. Pearce, and R. S. Cohen (eds.), *PSA 1974*. Dordrecht: Reidel, pp. 671-710. Reprinted in Sober, Elliott (ed.) (1984) *Conceptual Issues in Evolutionary Biology: An Anthology*, 1st. ed. Cambridge, MA: MIT Press, pp. 477-508.
- Wimsatt, William, C., (1980) *Reductionistic Research Strategies and their Biases in the Units of Selection Controversy*, in T. Nickles, éd., *Scientific Discovery*, v. II: *Case Studies*, Dordrecht: D. Reidel, pp. 213-259.
- Wimsatt, William (1986) *Forms of Aggregativity*, in A. Donagan, A. N. Perovich, Jr., and M. V. Wedin (eds.), *Human Nature and Natural Knowledge*, 259-291.

- Wimsatt, William, C. (1994) *The Ontology of Complex Systems: Levels of Organization, Perspectives, and Causal Thickets*, Canadian Journal of Philosophy, Supplementary Volume 20.
- Wimsatt, William, C. (1996) *Aggregativity: Heuristics for Finding Emergence*, Philosophy of Science Association, Vol. 64, Supplement, Proceedings of the 1996 Biennial Meetings of the Philosophy of Science Association, Part II: Symposia Papers, (Dec. 1997), p. S372-S384.
- Woodward, James (1989) *The Causal Mechanical Model of Explanation*, in *Scientific Explanation*, P. Kitcher et W. C. Salmon (ed.), Minnesota Studies in the Philosophy of Science, Vol. XIII.
- Woodward, James (2002) *What is a Mechanism? A Counterfactual Account*, Philosophy of Science (Supplement) 69: S366-S377.
- Woodward, James (2003) *Making things Happen*, A Theory of Causal Explanation, Oxford University Press.